



# **STIC Search Report**

## **Biotech-Chem Library**

**STIC Database Tracking Number:** 118820

**TO: Manjunath N Rao**  
**Location: REM-3B81&3C70**  
**Art Unit: 1652**  
**Wednesday, April 07, 2004**

**Case Serial Number: 09/246451**

**From: Noble Jarrell**  
**Location: Biotech-Chem Library**  
**Rem 1B71**  
**Phone: 272-2556**

**Noble.jarrell@uspto.gov**

### **Search Notes**

STIC-Biotech/ChemLib

118820

From: Rao, Manjunath N.  
Sent: Tuesday, April 06, 2004 10:54 AM  
To: STIC-Biotech/ChemLib  
Subject: Sequence search request for 09/246,451

From: Manjunath N. Rao  
Art Unit 1652, Room 3B-81  
Phone: 571-272-0939

Date: 4-6-04

Please search the following as soon as possible for application with serial number **09/246,451**

SEQ ID NOs. 2, 11, 12, 13 and 17 against all protein databases including issued patents database, published applications and pending application database and provide a print of all results.

If you have any questions please call me at the above phone number.

Thanks

Manjunath N. Rao  
Art Unit 1652  
Room 3B-81, Remsen Bldg.  
Phone: 272-0939

**Manjunath N. Rao**  
Art Unit 1652, Room 3B81,  
Mail Box in Room 3C70,  
Remsen Building, USPTO  
400, Dulany St.  
Alexandria, VA.  
Phone: 571-272-0939

Searching: Gunnell/Schreiber  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: \_\_\_\_\_  
Date Completed: 4/17/04  
Searcher Prep/Review: 5  
Clerical: \_\_\_\_\_  
Online time: 5

TYPE OF SEARCH:  
NA Sequences: \_\_\_\_\_  
AA Sequences: 5  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST (where applic.)  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: CompuGen  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 6, 2004, 18:53:57 ; Search time 34.974 Seconds  
(without alignments)  
3108.883 Million cell updates/sec

Title: US-09-246-451a-2  
Perfect score: 2180  
Sequence: 1 TTTTOSNANLAPLPHVPE.....IVSGVQALVWDPATTKAV 414

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1071772 seqs, 26263353 residues  
Total number of hits satisfying chosen parameters: 1071772

Minimum DB seq length: 0  
Maximum DB seq length: 2006000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA.\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US05\_PUBCOMB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US04\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US03\_PUBCOMB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/US02\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US01\_PUBCOMB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US00\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US05\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US04\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US03\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US02\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US01\_PUBCOMB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US00\_PUBCOMB.pep.\*

2red. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2180	100.0	414	15	US-10-453-104-2
2	2176	99.8	414	15	US-10-453-104-11
3	2169	99.5	414	15	US-10-453-104-12
4	2165	99.3	414	15	US-10-453-104-13
5	377.5	17.3	404	14	US-10-214-446-50
6	374.5	17.2	404	14	US-10-214-446-40
7	372	17.1	399	14	US-10-314-657-37
8	367.5	16.8	416	14	US-10-156-761-14997
9	363.5	16.7	416	9	US-09-861-289-39
10	363.5	16.7	416	9	US-09-860-846-39
11	363.5	16.7	416	10	US-09-988-384B-39
12	363.5	16.7	416	10	US-09-836-821-39
13	363.5	16.7	416	10	US-09-793-708-18
14	363.5	16.7	416	14	US-10-201-365-13
15	363.5	16.7	416	14	US-10-160-539-18

16 363.5 16.7 416 14 US-10-271-889-39 Sequence 39, Appl

17 359.5 16.5 411 14 US-10-156-761-8376 Sequence 376, Ap

18 356 16.3 393 14 US-10-156-761-9525 Sequence 9525, Ap

19 355 16.3 405 14 US-10-214-446-38 Sequence 38, Appl

20 354 16.2 425 14 US-10-214-446-20 Sequence 20, Appl

21 352.5 16.2 457 14 US-10-156-761-11073 Sequence 11073, A

22 348 16.3 418 12 US-10-389-647-559 Sequence 559, App

23 347.5 15.9 399 14 US-10-156-761-9914 Sequence 9914, Ap

24 339.5 15.6 388 14 US-10-156-761-13776 Sequence 13776, A

25 336.5 15.4 408 14 US-10-214-446-4 Sequence 4, Appl

26 333 15.3 409 15 US-10-458-201-12 Sequence 12, Appl

27 330 15.1 392 14 US-10-214-446-32 Sequence 32, Appl

28 328.5 15.1 406 14 US-10-214-446-2 Sequence 2, Appl

29 321 14.7 404 14 US-10-214-446-16 Sequence 16, Appl

30 320 14.7 401 14 US-10-156-761-8710 Sequence 8710, Ap

31 318 14.6 428 14 US-10-201-213-6 Sequence 6, Appl

32 317.5 14.6 396 14 US-10-156-761-14659 Sequence 14659, A

33 316.5 14.5 404 14 US-10-214-446-14 Sequence 14, Appl

34 315 14.4 412 14 US-10-214-446-36 Sequence 36, Appl

35 315 14.4 421 14 US-10-156-761-9703 Sequence 9703, Ap

36 314 14.4 418 14 US-10-214-446-22 Sequence 22, Appl

37 307 14.1 421 14 US-10-214-446-18 Sequence 18, Appl

38 304.5 14.0 400 14 US-10-314-657-62 Sequence 62, Appl

39 304.5 14.0 404 14 US-10-156-761-10431 Sequence 10431, A

40 300.5 13.8 415 14 US-10-214-446-46 Sequence 46, Appl

41 299 13.7 475 14 US-10-145-415-22 Sequence 22, Appl

42 298.5 13.7 402 14 US-10-205-032-8 Sequence 8, Appl

43 298 13.7 470 14 US-10-145-415-6 Sequence 6, Appl

44 296 13.6 429 14 US-10-145-415-14 Sequence 14, Appl

45 295 13.5 399 14 US-10-156-761-7959 Sequence 7959, Ap

## ALIGNMENTS

## RESULT 1

US-10-453-104-2  
; Sequence 2, Application US/10453104  
; Publication No. US20030207345A1  
; GENERAL INFORMATION:  
; APPLICANT: California Institute of Technology;  
; APPLICANT: Frances H. Arnold  
; APPLICANT: Hyun Joo  
; TITLE OF INVENTION: Oxygenase Enzymes and Screening Method  
; FILE REFERENCE: 4058/18827-US3  
; CURRENT APPLICATION NUMBER: US/10/453,104  
; CURRENT FILING DATE: 2003-06-02  
; PRIOR APPLICATION NUMBER: US 09/661,093  
; PRIOR FILING DATE: 2000-09-13  
; PRIOR APPLICATION NUMBER: US 09/246,451  
; PRIOR FILING DATE: 1999-02-09  
; PRIOR APPLICATION NUMBER: US 60/094,403  
; PRIOR FILING DATE: 1998-07-28  
; PRIOR APPLICATION NUMBER: US 60/106,840  
; PRIOR FILING DATE: 1998-11-03  
; PRIOR APPLICATION NUMBER: US 60/086,206  
; PRIOR FILING DATE: 1998-05-21  
; PRIOR APPLICATION NUMBER: US 60/106,834  
; PRIOR FILING DATE: 1998-11-03  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: FASTSEQ for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 414  
; TYPE: PRT  
; ORGANISM: P. Putida  
US-10-453-104-2

Query Match 100.0%; Score 2180; DB 15; Length 414;  
Best Local Similarity 100.0%; Pred. No. 1.3e-211;  
Matches 414; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTTTOSNANLAPLPHVPEHLVDFDYMNPNSLGSAGVQAWLQESNVDLWPTCNG 60

```

Db      1  TTETIQSNANLAPLPHPVPEHLVDFDMYNPNLSAGVQAEAWVLQESNVDPDLVWTRCNG 60
QY      61  GHWIATRGQIRAEYEDYRHFSSCEPFPREAGEAYDFITSMDDPPQRFALANQVVG 120
Db      61  GHWIATRGQIRAEYEDYRHFSSCEPFPREAGEAYDFITSMDDPPQRFALANQVVG 120
QY      121  MPVVDKLENRIQELACSLIESLRPOGQCNFTEDYAEPPPIRIFMLLAGLPEEDIPLHKYL 180
Db      121  MPVVDKLENRIQELACSLIESLRPOGQCNFTEDYAEPPPIRIFMLLAGLPEEDIPLHKYL 180
QY      181  TDQWTRPDGSMFAEAKAALYDYLPIIEQRORQKPGTDAISIVANGVNGRPITSDAKR 240
Db      181  TDQWTRPDGSMFAEAKAALYDYLPIIEQRORQKPGTDAISIVANGVNGRPITSDAKR 240
QY      241  MCGLLLVGGDLTVNFIUSFMEFLAKSPEHQELIERPERIPAAACEELLRRFSLVADGRI 300
Db      241  MCGLLLVGGDLTVNFIUSFMEFLAKSPEHQELIERPERIPAAACEELLRRFSLVADGRI 300
QY      301  LTSDEYFHGVLKGGDQILLPQMLSGLDERENACPMHVDFSRQKVSHTTFGHSHLCLGQ 360
Db      301  LTSDEYFHGVLKGGDQILLPQMLSGLDERENACPMHVDFSRQKVSHTTFGHSHLCLGQ 360
QY      361  HLARREIIVTLKWLTRIPDFSIAPGAQIOHKSGIVSGVQALPLVWDPATTKAV 414
Db      361  HLARREIIVTLKWLTRIPDFSIAPGAQIOHKSGIVSGVQALPLVWDPATTKAV 414

```

## RESULT 2

```

US-10-453-104-11
: Sequence 11, Application US/10453104
: Publication No. US20030207345A1
: GENERAL INFORMATION:
: APPLICANT: California Institute of Technology;
: APPLICANT: Frances H. Arnold
: APPLICANT: Hyun Joo
: TITLE OF INVENTION: Oxygenase Enzymes and Screening Method
: FILE REFERENCE: 4058/15827-US3
: CURRENT FILING DATE: 2003-06-02
: PRIOR FILING DATE: 2000-09-13
: PRIOR APPLICATION NUMBER: US 09/661,093
: PRIOR FILING DATE: 1999-02-09
: PRIOR APPLICATION NUMBER: US 60/094,403
: PRIOR FILING DATE: 1998-07-28
: PRIOR APPLICATION NUMBER: US 60/106,840
: PRIOR FILING DATE: 1998-11-03
: PRIOR APPLICATION NUMBER: US 60/086,206
: PRIOR FILING DATE: 1998-05-21
: PRIOR APPLICATION NUMBER: US 60/106,834
: NUMBER OF SEQ ID NOS: 19
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 11
: LENGTH: 414
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Mutant M7-4H
US-10-453-104-11

```

```

Query Match      99.8%; Score 2176; DB 15; Length 414;
Best Local Similarity 99.8%; Pred. No. 3.3e-211;
Matches 413; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1  TTETIQSNANLAPLPHPVPEHLVDFDMYNPNLSAGVQAEAWVLQESNVDPDLVWTRCNG 60
Db      1  TTETIQSNANLAPLPHPVPEHLVDFDMYNPNLSAGVQAEAWVLQESNVDPDLVWTRCNG 60
QY      61  GHWIATRGQIRAEYEDYRHFSSCEPFPREAGEAYDFITSMDDPPQRFALANQVVG 120
Db      61  GHWIATRGQIRAEYEDYRHFSSCEPFPREAGEAYDFITSMDDPPQRFALANQVVG 120

```

```

QY      121  MPVVDKLENRIQELACSLIESLRPOGQCNFTEDYAEPPPIRIFMLLAGLPEEDIPLHKYL 180
Db      121  MPVVDKLENRIQELACSLIESLRPOGQCNFTEDYAEPPPIRIFMLLAGLPEEDIPLHKYL 180
QY      181  TDQWTRPDGSMFAEAKAALYDYLPIIEQRORQKPGTDAISIVANGVNGRPITSDAKR 240
Db      181  TDQWTRPDGSMFAEAKAALYDYLPIIEQRORQKPGTDAISIVANGVNGRPITSDAKR 240
QY      241  MCGLLLVGGDLTVNFIUSFMEFLAKSPEHQELIERPERIPAAACEELLRRFSLVADGRI 300
Db      241  MCGLLLVGGDLTVNFIUSFMEFLAKSPEHQELIERPERIPAAACEELLRRFSLVADGRI 300
QY      301  LTSDEYFHGVLKGGDQILLPQMLSGLDERENACPMHVDFSRQKVSHTTFGHSHLCLGQ 360
Db      301  LTSDEYFHGVLKGGDQILLPQMLSGLDERENACPMHVDFSRQKVSHTTFGHSHLCLGQ 360
QY      361  HLARREIIVTLKWLTRIPDFSIAPGAQIOHKSGIVSGVQALPLVWDPATTKAV 414
Db      361  HLARREIIVTLKWLTRIPDFSIAPGAQIOHKSGIVSGVQALPLVWDPATTKAV 414

```

## RESULT 3

```

US-10-453-104-12
: Sequence 12, Application US/10453104
: Publication No. US20030207345A1
: GENERAL INFORMATION:
: APPLICANT: California Institute of Technology;
: APPLICANT: Frances H. Arnold
: APPLICANT: Hyun Joo
: TITLE OF INVENTION: Oxygenase Enzymes and Screening Method
: FILE REFERENCE: 4058/15827-US3
: CURRENT APPLICATION NUMBER: US/10/453,104
: PRIOR FILING DATE: 2003-06-02
: PRIOR APPLICATION NUMBER: US 09/661,093
: PRIOR FILING DATE: 2000-09-13
: PRIOR APPLICATION NUMBER: US 09/246,451
: PRIOR FILING DATE: 1999-02-09
: PRIOR APPLICATION NUMBER: US 60/094,403
: PRIOR FILING DATE: 1998-07-28
: PRIOR APPLICATION NUMBER: US 60/106,840
: PRIOR FILING DATE: 1998-11-03
: PRIOR APPLICATION NUMBER: US 60/086,206
: PRIOR FILING DATE: 1998-05-21
: PRIOR APPLICATION NUMBER: US 60/106,834
: NUMBER OF SEQ ID NOS: 19
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 12
: LENGTH: 414
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Mutant M7-6H
US-10-453-104-12

```

```

Query Match      99.5%; Score 2169; DB 15; Length 414;
Best Local Similarity 99.5%; Pred. No. 1.7e-210;
Matches 422; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1  TTETIQSNANLAPLPHPVPEHLVDFDMYNPNLSAGVQAEAWVLQESNVDPDLVWTRCNG 60
Db      1  TTETIQSNANLAPLPHPVPEHLVDFDMYNPNLSAGVQAEAWVLQESNVDPDLVWTRCNG 60
QY      61  GHWIATRGQIRAEYEDYRHFSSCEPFPREAGEAYDFITSMDDPPQRFALANQVVG 120
Db      61  GHWIATRGQIRAEYEDYRHFSSCEPFPREAGEAYDFITSMDDPPQRFALANQVVG 120
QY      121  MPVVDKLENRIQELACSLIESLRPOGQCNFTEDYAEPPPIRIFMLLAGLPEEDIPLHKYL 180
Db      121  MPVVDKLENRIQELACSLIESLRPOGQCNFTEDYAEPPPIRIFMLLAGLPEEDIPLHKYL 180
QY      181  TDQWTRPDGSMFAEAKAALYDYLPIIEQRORQKPGTDAISIVANGVNGRPITSDAKR 240

```

Db	181	TDQMRTPGSMTFABACGALXDYDLP	IIPIEQRRQKPGTDAISIVAGQVNGRPIITSD	EAKR	240
Qy	241	MCGLLVGGLGLOTVVNFISFSEMF	AKSPHQRQELTERPERIPAA	CEELRRRSLVADGRI	300
Db	241	MCGLLVGGLGLOTVVNFISFSEMF	AKSPHQRQELTERPERIPAA	CEELRRRSLVADGRI	300
Qy	301	LTSYEFHGVOLKKGQDILLPQML	SGLDERENACPMHVDPSQCKVSHHTTF	HGSHLCLGQ	360
Db	301	LTSYEFHGVOLKKGQDILLPQML	SGLDERENACPMHVDPSQCKVSHHTTF	HGSHLCLGQ	360
Qy	361	HLARREIIVTLKWLTRIPDSFIAP	GAQIOHKSGIVSGVQALPLVWD	PATTKAV	414
Db	361	HLARREIIVTLKWLTRIPDSFIAP	GAQIOHKSGIVSGVQALPLVWD	PATTKAV	414

```

RESULT 4
US-10-453-104-13
; Sequence 13, Application US/10453104
; Publication No. US2003207345A1
; GENERAL INFORMATION:
; APPLICANT: California Institute of Technology;
; APPLICANT: Frances H. Arnold
; APPLICANT: Hyun Joo
; TITLE OF INVENTION: Oxygenase Enzymes and Screening Method
; FILE REFERENCE: 4058/1E827-US3
; CURRENT APPLICATION NUMBER: US/10/453,104
; CURRENT FILING DATE: 2003-06-02
; PRIOR APPLICATION NUMBER: US 09/651,093
; PRIOR FILING DATE: 2000-09-13
; PRIOR APPLICATION NUMBER: US 09/246,451
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: US 60/094,403
; PRIOR FILING DATE: 1998-07-26
; PRIOR APPLICATION NUMBER: US 60/106,840
; PRIOR FILING DATE: 1998-11-03
; PRIOR APPLICATION NUMBER: US 60/086,206
; PRIOR FILING DATE: 1998-05-21
; PRIOR APPLICATION NUMBER: US 60/105,834
; PRIOR FILING DATE: 1998-11-03
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Fast-SEQ for Windows Version 3.0
; SEQ ID NO 13
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mutant M7-8H
US-10-453-104-13

```

QY	301	LTSDYEHGVLKKGDDILLPOMLSGLDERENACPMHVDVFRQKVSHSTFEGHSLCQG	360
Db	301	LTSDYEHGVLKKGDDILLPOMLSGLDERENACPMHVDVFRQKVSHSTFEGHSLCQG	360
QY	361	HLARREIVTLKWLTRIPDPSIAPGAQIQHKSGIVSGVALPLVWDPAATKAV	414
Db	361	HLARREIVTLKWLTRIPDPSIAPGAQIQHKSGIVSGVALPLVWDPAATKAV	414

```

RESULT 5
US-10-214-446-50
; Sequence 50, Application US10214446
; Publication No. US20030180742A1
; GENERAL INFORMATION:
; APPLICANT: Weiner, David
; APPLICANT: Burk, Mark J.
; APPLICANT: Kitchman, Tim
; APPLICANT: Pujol, Catherine
; APPLICANT: Richardson, Toby
; APPLICANT: Short, Jay M.
; TITLE OF INVENTION: P450 ENZYMES, NUCLEIC ACIDS ENCODING
; AND METHODS OF MAKING AND USING THEM
; FILE REFERENCE: 09010-500001
; CURRENT APPLICATION NUMBER: US/10/214,446
; CURRENT FILING DATE: 2002-08-05
; PRIOR APPLICATION NUMBER: US 60/309,497
; PRIOR FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 50
; LENGTH: 404
; TYPE: PRT
; ORGANISM: Bacterial
US-10-214-446-50

```

RESULT 6  
US-10-214-446-40  
Sequence 40, Application US/10214446  
Publication No. US20030180742A1  
GENERAL INFORMATION:  
APPLICANT: Weiner, David  
APPLICANT: Burk, Mark J.  
APPLICANT: Hitchman, Tim  
APPLICANT: Putol, Catherine



QY	174	IPHLKYLTDQWTRPGSWNTFAEAKALDYLIPIIEQSRQKPGTDAISIVANGVNGREI	233
Db	181	HDFFESQSRLLRGFGIAEVDARAQLDXYIALIDRKKKEPGDGLDDLIQEQLNRGV	240
QY	234	TSDAKMCGLLLVGGLDVTVNFSLFSMEFLAKSPHEQELIERPERIPAACEEILRRFS	293
Db	241	DRAEVSIALTLALLAGHETTANN:SLGFTTLRHPEQLAEURAEPLMPAAVEELL-RFL	299
QY	294	LVADG--RLTSDYEPFGVQLKGDQILLPQWLSGLDERENACFHVDFPSRCKVSHTFG	351
Db	300	SIADGLLRVATEDIEVAGTTIRADEGVVFATSVINRDAAGFAEPDADWHSARHIVAFG	359
QY	352	HGSHLCIGQLHAREIIVTLKWLTRPDFSI-APGAQIQHSG-IVSGVQALPLW	406
Db	360	FGHQCLQGNLRAEMETALGTLFERFGLRLAAPADEIPKPGFTIQGLMELPVTW	416

```

RECORD 9
US-09-861-289-39
; Sequence 39, Application US/09861289
; Patent No. US20020110897A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 602.438US1
; CURRENT APPLICATION NUMBER: US/09/861,289
; CURRENT FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 39
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-861-289-39

```

Query Match	16.7%;	Score	363.5;	DB	9;	Length	416;
Best Local Similarity	28.3%;	Pred. No.	1.le-27;				
Matches	97;	Conservative	61;	Mismatches	164;	Indels	21;
				Gaps	7;		
76	QY	EDYRHSSECP	PIPRAGEAYDF	IFPTSMDDP	PEQORQ	PAALANQV	GMVWVKLENRIQELA 135
		: : :	: : :	: : :	: : :	: : :	: : :
72	QW	KOWR--NST	PTPEAAALNHKLES	--DPRH	TRLRKLVA	REETMERVELL	RPVQEIv 127
		: : :	: : :	: : :	: : :	: : :	: : :
136	QY	CSLIESL--	RPOQCNCFT	EDYAEPP	PIRFIMLLAGL	PEEDI	PHLKVLTDMTREDGSMTF 193
		: : :	: : :	: : :	: : :	: : :	: : :
128	Qb	DGLVDAM	LAAPDGR	ELMESLAWP	PIITVISEL	GWFEEDRA	FRVWTDAFVDPDPAQA 187
		: : :	: : :	: : :	: : :	: : :	: : :
194	QY	AEAKELYDY	LIPIITEQR	KQKEGTDAIS	--IVANGVNGR	PIITSCAEAKRC	GLILLVGGLDT 252
		: : :	: : :	: : :	: : :	: : :	: : :
188	Qb	QTAXAEMSG	YILSRILDSK	RQGDGEDLL	SALVRTSD	EDGSLTSEEL	LGNHILLVAGHET 247
		: : :	: : :	: : :	: : :	: : :	: : :
253	QY	VWNFLSF	WEFLAKSP	EPHQELIER	PIPAACE	LLRRFSIVAGRI	ITTSYEF----- 307
		: : :	: : :	: : :	: : :	: : :	: : :
248	Qb	TVNLIA	NYALLSH	PCDGLAAL	RAMTLLD	GAVEMLR-----	YEGPVESATYRFPVFPV 302
		: : :	: : :	: : :	: : :	: : :	: : :
308	QY	--HG	VQLKKG	DQIILFC	MLSGD	ERENACPMH	VDPSROKVSHTTTS
		: : :	: : :	: : :	: : :	: : :	: : :
303	Qb	DLGD	TVI2AG	DTVLVWL	ADAHRT	PERFPDP	HRFDIRRTAGHLAF
		: : :	: : :	: : :	: : :	: : :	: : :
366	QY	EII	VTLEKWL	TRIDFS--	IAPGA	QIOHKS	GIVSGVQALPLVW 436
		: : :	: : :	: : :	: : :	: : :	: : :
363	Qb	EARTAV	ALLERC	EDLAD	VPGEIVW	YPNPMIR	GLKALPIRW 405
		: : :	: : :	: : :	: : :	: : :	: : :

ESULT 10  
S-09-860-846-39  
Sequence 39, Application US/09860846

```

: Patent No. US20020164742A1
:
: GENERAL INFORMATION:
:   APPLICANT: Sherman, D.H.
:   APPLICANT: Liu, H.
:   APPLICANT: Xue, Y.
:   APPLICANT: Zhao, L.
:
:   TITLE OF INVENTION: DNA encoding methymycin and pikromycin
:
:   FILE REFERENCE: 600.438US1
:
:   CURRENT APPLICATION NUMBER: US/09/860,846
:
:   CURRENT FILING DATE: 2001-05-18
:
:   PRIOR APPLICATION NUMBER: 09/105,537
:
:   PRIOR FILING DATE: 1998-06-26
:
:   NUMBER OF SEQ ID NOS: 43
:
:   SOFTWARE: FastSeq for Windows Version 3.0
:
:   SEQ ID NO 39
:
:   LENGTH: 416
:
:   TYPE: prnt
:
:   ORGANISM: Streptomyces venezuelae
:
:   US-09-860-846-39

```

[illegible]

```

RESULT 11
US-09-988-384B-39
; Sequence 39, Application US/09988384B
; Publication No. US20030073824A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.536US1
; CURRENT APPLICATION NUMBER: US/09/988,384B
; CURRENT FILING DATE: 2001-11-19
; PRIORITY APPLICATION NUMBER: PCT/US99/14398
; PRIORITY FILING DATE: 1999-06-25
; PRIORITY APPLICATION NUMBER: US 09/105,537
; PRIORITY FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 53
; SEQ ID NO 39
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-988-384B-39

```

```

Query Match      16.7%; Score 363.5; DS -0; Length 416;
Best Local Similarity 28.3%; Pred. No. 1.le-27;
Matches 97; Conservative 61; Mismatches 164; Indels 21; Gaps 7;

QY      76 EYRHFSCEQFIPREAGEAECFTPTSDPQRCFRALANQVGVMPVVKLENRIQELA 135
      :.: :.: :.: :.: :.: :.: :.: :.: :.: :.: :.: :.: :.: :.:
Db      72 KDWK--NSTTPTAEAAALNNMUES--DPPHRTKRLKLVAREFTMARVELLRVQEIIV 127

QY      136 CSLIESL--RFGQGCNFTEDYAEFPFPIRIFMLLAGLPEEDIPHLKYLTQDMTRDGGMTF 193
      :.: :.: :.: :.: :.: :.: :.: :.: :.: :.: :.: :.: :.:
Db      128 DGLVDMLAAEDGRADLMESLAWELPTVISELLGVPEPDEAAFRVWTDARFFDDPAQA 187

QY      194 AEAEALDYILPIEORRQPGGTDALS--IVANGVNGRPITSDEAKMCGLLLVGSELT 252
      :.: :.: :.: :.: :.: :.: :.: :.: :.: :.: :.: :.: :.:
Db      188 QTAAEMSGYLSRLIDSKRGODGDLISALVTSDEGSRLTSEELGMAHILLVAGHET 247

QY      253 VYNFLSFMELAKSPHQRQLIERPPIPAACELLERFLSVADGRILLTSDYEF----- 307
      :.: :.: :.: :.: :.: :.: :.: :.: :.: :.: :.: :.: :.:
Db      248 TVNLIANGMYALLSHPOLAALRADMTLLDGAVEEMLR-----YEGVESATYRFPPEVP 302

QY      308 --HGVLQKKGDIQLLPQLGLDERENACPMHVDFSRKVSHSTTFHGSHLCIGHIARR 365
      :.: :.: :.: :.: :.: :.: :.: :.: :.: :.: :.: :.: :.:
Db      303 DLDCGVIPAGDTVLVWLADAHRTPERFEDEHRFDIRDGTACHLAFGHIHCIGAPLARL 362

QY      366 EILVTLKWLTRIPDFS--TAPGAQIOHKSGIVSGVQALPLVW 406
      :.: :.: :.: :.: :.: :.: :.: :.: :.: :.: :.: :.: :.:
Db      363 EARIAVALLERCPLDALDVSFGLVWYVYPMINGLALPTRW 405

```

```

RESULT 12
JS-09-836-821-39
; Sequence 39, Application US/09836821
; Publication No. US20030087405A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/836,821
; CURRENT FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 39
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
JS-09-836-821-39

```

Query Match	-6.7%; Score 363.5; DB 1c; Length 416;
Best Local Similarity	Pred. No. 1.le-27;
Matches	97; Conservative 61; Mismatches 164; Idels 21; Gaps 7;
QY	76 EYRHFSSCPTIPREAGRAYOFIFTSMDPEQRQFRALANQVGVPPVDKLENRIQELA 135 : :::
Dd	72 KDWR--NSTPLTEAAALNNMLBS--DEPHRLRLKLVAKEFTWRRVELLRPRVQEIV 127 : :::
QY	136 CSLIESL--RPOQCNTFYDAPPIRFIMLAGLPEDDIPLHKYLTDQMTRPGSMFF 193 : :::
Dd	128 DELVDAMLAAPGDRADLMESLAWLPITVISLLGPEDRAAFRWTDFAVFDDPAQA 187 : :::
QY	194 ABAXDALYVLIPIIETOROKPQTDAIS--IVANGCWNERPITSDEAKMGCLLLVGGLDT 252 : :::
Dd	188 QTAMAEMSGYLSRLTOSRKGGQGEDLLSALVTSDSDGSLTSELLGNHAILLVAGHET 247 : :::
QY	253 VYNFISFMFLAKSPHEHQEIERPERIPAACEELLRRFRFSIVADGRILITSDFE----- 307 : :::
Dd	248 TWNLIAANGAYALLSHPDQLAALRADMT-LDGAVEMLR-----YEGPVESATYRFPVEPV 302 : :::
QY	308 --HGVLKKGDQILLPOMLSGLDERENACPMIEDFSROKVSHTTTFGHSHLCLOGLARR 365 : :::

```

Db      303  DLDGTIVAGTGLVVLADAHRTPRFRPDRFRD-RRDTAGHLAFGHGHIHFCIGAPLRL 362
QY      366  EIIIVTLKWLTRIPDFS--IAPGAQIOHKSIVSGVQALPLVW 406
Db      363  EARIAVRAILLERCPLALDVSPGELVWYPNPMIRGLKALPIRW 405

RESULT 13
US-09-793-708-18
: Sequence 18, Application US/09793708
: Publication No. US20030104597A1
: GENERAL INFORMATION:
: APPLICANT: ASHLEY, Gary
: APPLICANT: BETLACH, Melanie C.
: APPLICANT: BETLACH, Mary C.
: APPLICANT: McDANIEL, Robert
: APPLICANT: TANG, Li
: TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
: CURRENT APPLICATION NUMBER: 300622002121
: CURRENT FILING DATE: 2001-02-22
: PRIOR APPLICATION NUMBER: US 09/657,440
: PRIOR FILING DATE: 2000-09-07
: PRIOR APPLICATION NUMBER: US 09/320,878
: PRIOR FILING DATE: 1999-05-27
: PRIOR APPLICATION NUMBER: US 09/320,878
: PRIOR FILING DATE: 1999-05-27

```

```

; PRIOR APPLICATION NUMBER: US 09/073,538
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: US 08/846,247
; PRIOR FILING DATE: 1997-04-30
; PRIOR APPLICATION NUMBER: US 60/134,99C
; PRIOR FILING DATE: 1999-05-20
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-793-708-18

Query Match      16.7%; Score 363.5; DB 10; Length 416;
Best local Similarity 28.3%; Pred.No.1.le-27;
Matches          97; Conservative 61; Mismatches 164; Indels 21; Gaps 7;

QY       76 EDVRFHSSECFPIPRAGEAYDIPITSMDEPQRGFALANQVGVPTVKLENRIQELA 135
         :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:|
Db        72 KDMR--NSTTPLTEAAGLNNHMLCS--DPRETRLRKLVAREFTMRVRVELLPYQEIV 127

```

```

128 DGUVDAWLAAPDGRADIMESLAWPLFITWISSELLGVPEPDRAAFRVYVMTDAFVPDDPAQA 187
194 AEAKEALYDYLPIIEQRQKPGTDAIS--IVANGQVNGRPITSDAEKRMCGILLVGLDIT 252
188 QTAAEMSGYLSLIDSKGQGEDULLSALVRTSDGDSGLTSEELLGMAHILLVAGHET 247
253 VVNFLSFSMEFLAKSPEHQELIERPERIPAAECELLRRFLSVADGRILITSDYEF----- 307
248 TVNLIANGCVALLSHPDQLAALRADMTLLDGAVEMLR-----YEGPVSATYRFPPEPV 302
308 --HGVLKKGDDQLLPLMGLSGIDDERENACPMGVDFSRQKVSHTTFGHSHLCIGQHLARR 365
303 DLDGTVIPAGDVTVLVLAHAHTPFRFPDPEHRFDIRRDTAGHLAFGHGTHFCIGAPLAKL 362
366 EIIVTLKEMLTRIPDPS--IAPGAQIQKSGIVSGVQCALPLYW 406
363 EARIAVRALLERCPDIALDVSPGELVWYDNPWIRGKALPIRW 405

```



	Query Match	16.7%	Score 363.5;	DB 14;	Length 416;
	Best Local Similarity	28.3%;	Pred. No. 11e-27;		
	Matches	97;	Conservative 61;	Mismatches 164;	Indels 21; Gaps 7;
QY	76	EDYRHSSCCFFIPREAGRAYDIFPTSMDDPEQQRFRALANQVGVGVVVKLENRIQELA	135		
Db	72	KDWR--NSTTETLAEAAALHNHMLSS--DPPHTRIRLKLVARFETMRVVELLPVQVEIV	127		
QY	136	CSLIESL--RPGQCNTEDYAEPPFIRIFMLLAGHPEEDIPHLKYLTIDQMTRPDGSMTF	193		
Db	128	DGLVDAMLAAAPDGRADLMESLAWPFEITVISLGVPEPDRAAFRVTWDAFVFPDDPAQA	187		
QY	194	AAAKALVLYLPITIEQRQKFGCTDAIS-IVANGQVNGRPITISDEAKRMCGILLVGGLOT	252		

QY	76	EDYHRFSSECCPFTPRAGEAYDFTPTSMDEPQEQFRALANQVGMPPVVDKXERIOELA	135
Db	72	KDWR--NSTTPEAEAAAHNMLES--DDPHTRLRKLVAREFTMRVELLRERVQIIV	127
QY	136	CSLIESL--REQQCNTEDYAPFPFTRIFMLLAGLFEEDIPHKLYLTQDMTRPDGSMTF	193
Db	128	DGLVDAMLAPDGRADIMESLAWPLRTIVISELLGVPEPORAFAKWWTDAVFVDDPAQA	187
QY	194	AEAKALYDIPIIEORROKEGTDAIS--IVANGQVNGRPTITSEARMCGLLIVGGLDT	252
Db	188	QTAAWMSGYLSRLIISKQGDGEDLSALVTSDEGSRLTSEELGMAFILLVAGHET	247
QY	253	VNFTSFMFLAKSPHROELIERPERIPAAACELLRRFSLVADGRILTSDYEF-----	307
Db	248	TVNLIANGVYALLSHEDQLAALRADMTLLDGAVEMLR-----YEGVSESATYFPFVWPV	302
QY	308	--HGVLKKGDDQILLFCQMLSGLDERENACPMHVDPSQKYSHTTFFGHGSHLCQHLARR	365
Db	303	ELEGTIVIPAGDTVLVVLADAHRTPFEPDPHREDRTAGHLAFHGCHFCIGCAPLARL	362
QY	366	EIIVTLKWLTRIPDFS--IAPCAQIOHKSGVSGVQALPLWV	406
Db	363	EARIARALLERCPDALVDSPGLWYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV	405

Search completed: April. 6, 2004, 19:14:29  
Job time : 36.974 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 6, 2004, 18:49:16 ; Search time 14.9588 Seconds  
(without alignments)  
1428.803 Million cell updates/sec

Title: US-09-246-451A-2

Perfect score: 2180

Sequence: 1 TTEIQQNANLAPPPHVPF.....IVSGVQALPFWDPATTYKAV 414

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA.\*

1: /cgn2\_6/ptodata/2/iaa/5A COMB.pap.\*

2: /cgn2\_6/ptodata/2/iaa/5B COMB.pap.\*

3: /cgn2\_6/ptodata/2/iaa/6A COMB.pap.\*

4: /cgn2\_6/ptodata/2/iaa/6B COMB.pap.\*

5: /cgn2\_6/ptodata/2/iaa/PCTUS COMB.pap.\*

6: /cgn2\_6/ptodata/2/iaa/backfiles.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	363.5	16.7	416	3	US-09-320-878-18
2	363.5	16.7	416	3	US-09-105-537-39
3	363.5	16.7	416	4	US-09-141-908-13
4	363.5	16.7	416	4	US-09-657-408-18
5	347	15.9	437	4	US-09-252-991A-17836
6	340	15.6	403	6	5212296-9
7	333	15.3	409	3	US-09-385-028-12
8	332	15.3	409	4	US-09-726-614-12
9	333	15.3	409	4	US-09-385-040-12
10	330	15.1	406	6	5212296-6
11	324.5	14.9	399	4	US-08-765-907A-10
12	316	14.5	412	1	US-08-102-863-11
13	316	14.5	412	5	PCT-US92-10885-11
14	271.5	12.5	419	3	US-09-335-409-8
15	271.5	12.5	419	3	US-09-413-814-71
16	271.5	12.5	419	4	US-09-568-102-8
17	271.5	12.5	419	4	US-09-567-969-8
18	271.5	12.5	419	4	US-09-568-480-8
19	271.5	12.5	419	4	US-09-568-486-8
20	271.5	12.5	419	4	US-09-568-472-8
21	271.5	12.5	419	4	US-09-567-899-8
22	261	12.0	395	4	US-09-266-965-129
23	233	10.7	468	4	US-09-252-991A-32437
24	169.5	7.8	189	4	US-09-679-279-20
25	161	7.4	422	2	US-09-096-982-5
26	161	7.4	422	2	US-08-653-650A-5
27	161	7.4	474	2	US-09-096-982-8

#### ALIGNMENTS

##### RESULT 1

US-09-320-878--8

; Sequence 18, Application US/09320878A

; Patent No. 6117659

; GENERAL INFORMATION:

; APPLICANT: ASHLEY, Gary

; APPLICANT: BETLACH, Melanie C.

; APPLICANT: BETLACH, Mary C.

; APPLICANT: McDANIEL, Robert

; APPLICANT: TANG, Li

; TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE

; CURRENT APPLICATION NUMBER: US/09/320,878A

; CURRENT FILING DATE: 1999-05-27

; EARLIER APPLICATION NUMBER: CIP OF 09/141,908

; EARLIER FILING DATE: 1998-08-28

; EARLIER APPLICATION NUMBER: CIP OF 09/073,538

; EARLIER FILING DATE: 1998-05-06

; EARLIER APPLICATION NUMBER: CIP OF 08/846,247

; EARLIER FILING DATE: 1997-04-30

; EARLIER APPLICATION NUMBER: 60/119,139

; EARLIER FILING DATE: 1999-02-08

; EARLIER APPLICATION NUMBER: 60/100,880

; EARLIER APPLICATION NUMBER: 60/087,080

; EARLIER FILING DATE: 1998-05-28

; NUMBER OF SEQ ID NOS: 34

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 18

; LENGTH: 416

; TYPE: PRT

; ORGANISM: Streptomyces venezuelae

US-09-320-878--8

Query Match

Best Local Similarity 16.7%; Score 163.5; DB 3; Length 416;

Matches 9; Conservative 61; Mismatches 164; Indels 21; Gaps 7;

QY	76	EDYHSESCEFFIPREAGEAYDFIPTSMPPQRFALANQVGVWVQKLENRIOELA	135
Db	72	KDWR--NSTPLTEAEALNHNLES--DPFRTRUKLVAREFTMRVLRPRVQEV	127
QY	136	CSLIESI--RPQCCHFTEDYAPFFPIRIFMLAGUPEEDIPHLKYLTQMTDPGSMTF	193
Db	128	DGLVDLAAPDGRADIMESLAWPLPITVISSELLGVPEPDRAAFRWVTDFAVFPDDPAQA	187
QY	194	AEAKEALYDILPI--FQRROKPCPTDAIS--IVANGVNGRPTISDEAKRMCGLLIVGLDT	252
Db	188	QTAAEMSGYLSRLIDSKRGQDGEDLLSALVRTSDEGSRLTSEELGMAHILLVAGHET	247

Sequence 8, Appli  
Sequence 9, Appli  
Sequence 9, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 5, Appli  
Sequence 2, Appli  
Sequence 4, Appli  
Sequence 2, Appli  
Sequence 7, Appli  
Sequence 7, Appli  
Sequence 22, Appli  
Sequence 2, Appli  
Sequence 24, Appli  
Sequence 6, Appli  
Sequence 7, Appli

QY 253 VVNFSLFMSFLAKSPHQRQLIERPERIPAAACEILLRFRSLVADGRILTSDYEF----- 307  
Db 248 TVNLIANGMYALLSHPOQALRADMTLLDGAVEMLR-----YEGPVESATYRFPPEVP 302  
QY 308 --HGVLKKGQDQILLPQMLSGLDRENACPMHVSFRQKVSHTTFGSHLCLGQHLARR 365  
Db 303 DLDTGTVIPAGDTVLVVLADAHRTPEFDPHFRDIRDTAGHLAFGHIHFCIGAPLARR 362  
QY 366 EIIIVTKEWLTIRIPDFS--IAPGAQIQHKSGLVSGVQALPLVW 406  
Db 363 EARIARALLERCPDLALDVSPGELVWYFNPMPMIRGLKALPIRW 405

RESULT 2

US-09-105-537-39  
; Sequence 39, Application US/09105537A  
; Patent No. 6265202  
; GENERAL INFORMATION:  
; APPLICANT: Sherman, P. H.  
; APPLICANT: Liu, H.  
; APPLICANT: Xue, Y.  
; APPLICANT: Zhao, L.  
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin  
; FILE REFERENCE: 600.438US1  
; CURRENT APPLICATION NUMBER: US/09/105,537A  
; CURRENT FILING DATE: 1998-06-26  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 39  
; LENGTH: 416  
; TYPE: PRT  
; ORGANISM: Streptomyces venezuelae  
US-09-105-537-39

Query Match 16.7%; Score 363.5; DB 3; Length 416;  
Best Local Similarity 28.3%; Pred. No. 3.4e-29;  
Matches 97; Conservative 61; Mismatches 164; Indels 21; Gaps 7;

QY 76 EDYRHSFSECPPIPREAGEAYDPTSMDDPEQRFALANQVGVMPVVDKLENRIQELA 135  
Db 72 KQWR--NSTTPLTEAEALNHNMLE--DPPRHLRLKLVAREFTMRVRELLPRVQEI 127  
QY 136 CSLIESL--RPOQCNTEDYAEPPFIRIFMLAGLPEEDIPHLKYLTDQTRPDGSMTF 193  
Db 128 DGLVDAMLAAPDGRADLMESLAWPLPTITVISELLGVPEPDRRAFRVMTDAFVFPDDPAQ 187  
QY 194 AEAKEALYDLIPIEOROKPCTDAIS--IVANGQVNGRPITSDAKRMCGILLVGLDT 252  
Db 188 QTAMAEKSGYLSRLDSKRGQDGLLSALVTSDEGSLTSELLGMAHILLVAGHET 247  
QY 253 VVNFSLFMSFLAKSPHQRQLIERPERIPAAACEILLRFRSLVADGRILTSDYEF----- 307  
Db 248 TVNLIANGMYALLSHPOQALRADMTLLDGAVEMLR-----YEGPVESATYRFPPEVP 302  
QY 308 --HGVLKKGQDQILLPQMLSGLDRENACPMHVSFRQKVSHTTFGSHLCLGQHLARR 365  
Db 303 DLDTGTVIPAGDTVLVVLADAHRTPEFDPHFRDIRDTAGHLAFGHIHFCIGAPLARR 362  
QY 366 EIIIVTKEWLTIRIPDFS--IAPGAQIQHKSGLVSGVQALPLVW 406  
Db 363 EARIARALLERCPDLALDVSPGELVWYFNPMPMIRGLKALPIRW 405

RESULT 3

US-09-141-908-13  
; Sequence 13, Application US/09141508  
; Patent No. 6503741  
; GENERAL INFORMATION:  
; APPLICANT: Ashley, Gary  
; APPLICANT: Betlach, Melanie C.  
; APPLICANT: Betlach, Mary  
; APPLICANT: McDaniel, Robert  
; APPLICANT: Tang, Li

; TITLE OF INVENTION: Combinatorial Polyketide Libraries Produced Using a  
; TITLE OF INVENTION: Modular PKS Gene Cluster as Scaffold  
; FILE REFERENCE: 300622002100  
; CURRENT APPLICATION NUMBER: US/09/141,908  
; CURRENT FILING DATE: 1998-08-28  
; EARLIER APPLICATION NUMBER: CIP OF 09/373,538  
; EARLIER FILING DATE: 1998-05-06  
; EARLIER APPLICATION NUMBER: CIP OF 08/846,247  
; EARLIER FILING DATE: 1997-04-30  
; EARLIER APPLICATION NUMBER: PROV. 60/076,919  
; EARLIER FILING DATE: 1998-03-05  
; EARLIER APPLICATION NUMBER: PROV. 60/087,080  
; EARLIER FILING DATE: 1998-05-28  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 13  
; LENGTH: 416  
; TYPE: PRT  
; ORGANISM: Streptomyces venezuelae  
US-09-141-908-13

Query Match 16.7%; Score 363.5; DB 4; Length 416;

Best Local Similarity 28.3%; Pred. No. 3.4e-29;  
Matches 97; Conservative 61; Mismatches 164; Indels 21; Gaps 7;  
QY 76 EDYRHSFSECPPIPREAGEAYDPTSMDDPEQRFALANQVGVMPVVDKLENRIQELA 135  
Db 72 KQWR--NSTTPLTEAEALNHNMLE--DPPRHLRLKLVAREFTMRVRELLPRVQEI 127  
QY 136 CSLIESL--RPOQCNTEDYAEPPFIRIFMLAGLPEEDIPHLKYLTDQTRPDGSMTF 193  
Db 128 DGLVDAMLAAPDGRADLMESLAWPLPTITVISELLGVPEPDRRAFRVMTDAFVFPDDPAQ 187  
QY 194 AEAKEALYDLIPIEOROKPCTDAIS--IVANGQVNGRPITSDAKRMCGILLVGLDT 252  
Db 188 QTAMAEKSGYLSRLDSKRGQDGLLSALVTSDEGSLTSELLGMAHILLVAGHET 247  
QY 253 VVNFSLFMSFLAKSPHQRQLIERPERIPAAACEILLRFRSLVADGRILTSDYEF----- 307  
Db 248 TVNLIANGMYALLSHPOQALRADMTLLDGAVEMLR-----YEGPVESATYRFPPEVP 302  
QY 308 --HGVLKKGQDQILLPQMLSGLDRENACPMHVSFRQKVSHTTFGSHLCLGQHLARR 365  
Db 303 DLDTGTVIPAGDTVLVVLADAHRTPEFDPHFRDIRDTAGHLAFGHIHFCIGAPLARR 362  
QY 366 EIIIVTKEWLTIRIPDFS--IAPGAQIQHKSGLVSGVQALPLVW 406  
Db 363 EARIARALLERCPDLALDVSPGELVWYFNPMPMIRGLKALPIRW 405

RESULT 4

US-09-657-440-18  
; Sequence 18, Application US/09657440  
; Patent No. 6509455  
; GENERAL INFORMATION:  
; APPLICANT: Ashley, Gary  
; APPLICANT: Betlach, Melanie C.  
; APPLICANT: Betlach, Mary C.  
; APPLICANT: McDaniel, Robert  
; APPLICANT: Tang, Li  
; TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE  
; FILE REFERENCE: 300622002120  
; CURRENT APPLICATION NUMBER: US/09/657,440  
; CURRENT FILING DATE: 2000-09-07  
; PRIOR APPLICATION NUMBER: 09/320,878  
; PRIOR FILING DATE: 1999-05-27  
; PRIOR APPLICATION NUMBER: CIP OF 09/141,908  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 34  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 18  
; LENGTH: 416  
; TYPE: PRT





Db 175 EDRSAVLDRGYTPE---QVAKARDELGYLRELVEERIENPGTDLISRLVIDQVRPGHL 231  
QY 234 TSDEAKRMCGLLVGLGLDVTNNFLSPSMEFLAKSEHRCQELTERPERIPACBELLRRFS 293  
Db 232 RVEEMVPCRLLLVAGHGTTTSQASLSLSLTDPDELAGRLTEDPALLPKAVEELLRRFS 291  
QY 294 LVADG--RLTSDYEFHGVOLKKGQOILLPQMLSGLDRENAACPMHVDPSROKVSHTTFFG 351  
Db 292 IVQNGLABAAVEDVQ--LDVLRAGEGVVLSAGNRDETVPDPORVDVDRDARRHLAFG 351  
QY 352 HGSHLCLGQHLAR---REIIVTLKEMLRIPDFSTA-PCAQIOHKSGIVS-GVQALPLVM 406  
Db 352 HGMEQCLGQHLARVELEELAAVLKWM---PGARLAVPFEELDFRHEVSSYGLGALPVTW 408  
QY 407 D 407  
Db 409 Z 409

RESULT 9  
US-09-385-040-12  
; Sequence 12, Application US/09385040  
; Patent No. 6589775  
; GENERAL INFORMATION:  
; APPLICANT: Jensen, Susan E  
; APPLICANT: Aicoo, Kwamena A  
; APPLICANT: Paradkar, Ashish S  
; TITLE OF INVENTION: DNA SEQUENCE ENCODING ENZYMES OF CLAVULANIC ACID  
; FILE REFERENCE: 09/385,040  
; CURRENT APPLICATION NUMBER: US/09/385,040  
; PRIOR FILING DATE: 1999-08-30  
; PRIOR APPLICATION NUMBER: US 08/790,462  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 12  
; TYPE: PRT  
; ORGANISM: Streptomyces clavuligerus  
US-09-385-040-12

Query Match 15.3%; Score 333; DB 4; Length 409;  
Best Local Similarity 29.0%; Pred. No. 5e-26;  
Matches 122; Conservative 60; Mismatches 185; Indels 54; Gaps 18;

QY 16 PHVPEHLWFDMDYNPNSLSAGVQERAWVLCSNVDLVWTRCNGH-WIATRGQLIREA 74  
Db 14 PAYPHERVCPVD---PPQLAGLRSSQKAAASRT-----LW---DGSQVMLVTSHAGARAV 62  
QY 75 YEDYRHS-SECPFIP-----REAGEAVDFIPTSDPPQRCFRA-----LANOV 118  
Db 63 LGDRRTAVTSAPGFMRLTRTSLVREANPESASF--RMDDPQHSRLRSMWTRDFLARRA 120  
QY 119 VGM-PWVOKLENRIQELACSLIESLRPQOCNFTEDYAEFPPIRIPMLLAGLFEEDIPHL 177  
Db 121 EALRPVAVREL---LDEILGLGVKGRP---VDLVAGTTPVPSRVITLLFGADDDRREFI 174  
QY 178 K----YLTDMQTRPGSMTFAEAKALDYLIPIEORQKPGTDAISIVANGQNGRPI 233  
Db 175 EDRSAVLDRGYTPE---QVAKARDELGYLRELVEERIENPGTDLISRLVIDQVRPGHL 231  
QY 234 TSDEAKRMCGLLVGLGLDVTNNFLSPSMEFLAKSEHRCQELTERPERIPACBELLRRFS 293  
Db 232 RVEEMVPCRLLLVAGHGTTTSQASLSLSLTDPDELAGRLTEDPALLPKAVEELLRRFS 291  
QY 294 LVADG--RLTSDYEFHGVOLKKGQOILLPQMLSGLDRENAACPMHVDPSROKVSHTTFFG 351  
Db 292 IVQNGLABAAVEDVQ--LDVLRAGEGVVLSAGNRDETVPDPORVDVDRDARRHLAFG 351  
QY 352 HGSHLCLGQHLAR---REIIVTLKEMLRIPDFSTA-PCAQIOHKSGIVS-GVQALPLVM 406  
Db 352 HGMEQCLGQHLARVELEELAAVLKWM---PGARLAVPFEELDFRHEVSSYGLGALPVTW 408

QY 407 D 407  
Db 409 Z 409

RESULT 10  
5212296-6  
; Patent No. 5212296  
; APPLICANT: DEAN, CAROLINE; HARDER, PATRICIA A.; LETO, KENNETH  
; J.; O'KEEFE, DANIEL P.; OMER, CHARLES A.; ROMESSER, JAMES A.  
; ZEPPERMAN, JAMES M.  
; TITLE OF INVENTION: EXPRESSION OF HERBICIDE METABOLIZING  
; CYTOCHROMES  
; NUMBER OF SEQUENCES: 19  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/569,781  
; FILING DATE: 23-AUG-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 464,499  
; FILING DATE: 12-JAN-1990  
; APPLICATION NUMBER: 425,605  
; FILING DATE: 11-SEP-1989  
; SEQ ID NO: 6;  
; LENGTH: 406  
5212296-6

Query Match 15.1%; Score 330; DB 6; Length 406;  
Best Local Similarity 26.3%; Pred. No. 1e-25;  
Matches 104; Conservative 74; Mismatches 193; Indels 24; Gaps 11;

QY 31 PSNLSAGVQ--EAWAVLOESNVPLVTRCNGH-WIATRGQLIREAYEDYR----- 79  
Db 17 PSNRSCPQLPDGVAQLRDTGPHLRVTLYDGRQAVVTKHEAARKLGLDPLSSNRD 76  
QY 80 HFSECECFIP--REAGEAVDFIPTSDPPQRCFALANQVGMVPMVDKLENRIQELACS 137  
Db 77 RFATSPRFRAVRESPOAF-----IGLDPPEHGTTRMTTSEFTVKRIKGMPEVEVHVG 132  
QY 138 LIESLRPQG-QCNFTEDYAEFPPIRIPMLLAGLFEEDIPHLKYLTDMQTRPGSMTFAEA 196  
Db 133 FLDEMLAAGTADIVSQFALPVPVSWVICRLGVYADHEFFQDASKRLVQSTDAQSALTA 192  
QY 197 KEALYDYLPIIQRRQKPGTDAI-SIVANGQNGRPIITSDAEKEMOGLLIVAGGLDVTWN 255  
Db 193 RNDLAGYLDGLITQFQTEPGAGLVGALVADQLANGE-IDREELISTAMLLIAGHETTAS 251  
QY 256 FLFSFMEFLAKSPEHRCQELTERPERIPACBELLRRFSL--VADGRILTSDEYEFHGVOLK 313  
Db 252 MTSLSVITLLDHPQYAAALRADSLVPGAVEELLRYLAADIAGGRVATADIEVEGLIR 311  
QY 314 KGOQILLPQMLSGLDRENAACPMHVDPSROKVSHTTFFGHSHLCLGQHLARREIIVTLKE 373  
Db 312 AGEVTVVANSIANDGTGVYEDPDALDIHRSARHHLAFGFGVHQCGLGONLARLEVLINA 371  
QY 374 WLTRIPDFSTA-PCAQIOHKSG-IVSGVQALPLVM 406  
Db 372 LMDRVETLRAVPVEQLVLRPGTTIQGVNELPVTW 406

RESULT 11  
US-08-765-907A-10  
; Sequence 10, Application US/08765907A  
; Patent No. 6352839  
; GENERAL INFORMATION:  
; APPLICANT: BLANC, Veronique  
; APPLICANT: THIBAUT, Denis  
; APPLICANT: BAKAS-JACQUES, Nathalie  
; APPLICANT: BLANCHE, Francis  
; APPLICANT: COUZET, Joel  
; APPLICANT: BARRIERE, Jean-Claude  
; APPLICANT: DEBUSSCHE, Laurent  
; APPLICANT: FMECHON, Alain

```

; APPLICANT: PARIS, Jean-Marc
; TITLE OF INVENTION: Streptogramins And Method For Preparing Same By
; FILE OF INVENTION: Mutasynthesis
; FILE REFERENCE: Streptomycin genes
; CURRENT APPLICATION NUMBER: US/08/765,907A
; CURRENT FILING DATE: 1997-03-20
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Streptomyces pristinaespiralis
US-08-765-907A-10

Query Match      : 4.9%; Score 124.5; DB 4; Length 399;
Best Local Similarity 29.0%; Pred. No. 3.7e-25;
Matches 106; Conservative 59; Mismatches 159; Indels 41; Gaps 12;

QY   74 AYEDYRH-----ESSECFPTPREAGEAYDFTPSMCPPEQOPFALANOVVGM 121
Db   36 AFHYFRHADVLTVASDPGVYSQSLSRLRPGSCALSEQILSVIDPPMHTLRIVSQAFTP 95

QY   122 EVDVKLENRIQLACSLIESLRPGQC-KFTEDVAEPPPIRFMLLAGLPEDPDHLKYL 180
Db   96 RTVADLEPRVTELAGQLDAV--DGDTFDLVAQFAFLPVIVIAELLGVVPADRTLPSFM 153

QY   181 TDQWTR-----PDGSMTEFAEAKELYDIPLIETQRQKPGETAISIVA 224
Db   154 SDRMLQGVADPADMQFDADAEDYQRLVKPMEAHAYLHDEVTDRAAPANDLLISALV 213

QY   225 NGCVNGRPITSDBAKRCGLILLVGLTIVNFLTFSMEFLAKSPEHQELIERPER--IP 282
Db   214 AARVEGERLTDEQIVEFGALLMHGWHSVTMLGNLTVLCLKDHP--RAEAAARADRSLLP 271

QY   283 AACEECLR-RFSLVADGRILTSYEFHFVGVLKKGDQILLPQMLS-GLDERENACPMEWDF 340
Db   272 ALIEEVLRIRPPIVMARVTTTKDTVLAGTTIPAG-RMVVFSLLSANHDEQVFTDPDEL 330

QY   341 SRQKVSHITFGHSHLCGCEHARRREIIVTLKEWLTRIPDFSAPGAQIQ-HKSIGVSGV 399
Db   331 ARRG-RQAIFGHGIHYCLGAPLARLEGRIALEALFORFPDPSPTDGAKLYHRDGLF-GV 388

QY   400 QALPL 404
Db   389 KNPLP 393

```

RESULT 12  
US-08-102-863-11  
; Sequence 11, Application US/08102863  
; Patent No. 5465590  
; GENERAL INFORMATION:  
; APPLICANT: SAKIASLANI, SIMA  
; TITLE OF INVENTION: CONSTITUTIVE  
; TITLE OF INVENTION: EXPRESSION OF P450SOY  
; TITLE OF INVENTION: AND FERREDOXIN-SOY IN  
; TITLE OF INVENTION: STREPTOMYCES  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: E. I. DU PONT DE NEMOURS  
; ADDRESSEE: AND COMPANY  
; STREET: 1007 MARKET STREET  
; CITY: WILMINGTON  
; STATE: DELAWARE  
; COUNTRY: USA  
; ZIP: 19898  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0,  
; SOFTWARE: Version #1.25

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/102,863
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/807,001
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: GALLEGOS, R. THOMAS
REGISTRATION NUMBER: 32,692
REFERENCE/DOCKET NUMBER: CR-9000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 302-892-7342
TELEFAX: 302-892-7949
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 412 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-102-863-11

Query Match      14.5%; Score 316; DB 1; Length 412;
Best Local Similarity 27.2%; Pred. No. 3e-24;
Matches      84; Conservative      56; Mismatches 163; Indels      6; Gaps      5;

QY      103 MDPEQQQPALANVGVMPVKLENRIQLACSLLESRPQG-QCNFTEDYAEPPIR 161
Db      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      105 VDDEHNTQRMLIPTFSVRKIGALRPRIQETVDRLDAMERQPPAELVSFAFPVSM 164
Db      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      162 IFMLLAGLPEDIPHLKYLTDOMTRPDGSMTFARAEKALDYILPIIEQRRKPGTDALS 221
Db      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      165 VICALLGPYADIAFAFEERSQRLIRGGADDVNRAKDDELBEYLICALIDRKRAEFGGLLD 224
Db      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      222 IVANGVNGRPITTSDEAKRMCGLLLVGLPTVVNF-SFSMEFLAKSPHRQELIERPERI 281
Db      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      225 ELIHDPDHPDQREQLVAFAVILLIAGHETTANNISLGFTELLSHPEQLAALRAGGTST 284
Db      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      282 PAACEELLRFESLVADG--RILTSDYEFGHVQLKGDDOILLPQMLSGLDERENACPMVD 339
Db      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      285 AVTVSELL-RFLSIAGLQRLATEDMEVDGATIRKGEVGVFSTSLNRRDADVFPRAETILD 343
Db      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      340 FSRQKYSHTFFRGSHLCGLQHILAREIIIVTKEWLTRIPDFSIA-PGAQIQHKS-G-IVS 397
Db      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      344 WDRPARHHLAFGVGVCQCLGNLARAEIDIAMTELFERLPGLAVPAHEIRHKPGDTTQ 403
Db      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      398 GVQAQLPVW 406
Db      ||:|||:
QY      404 GLLDLPWA 412
Db      ||:|||:

RESULT 13
PCT-US92-10885-11
Sequence 11, Application PC/TUS9210885
GENERAL INFORMATION:
APPLICANT: SARIASLANI, SIMA
TITLE OF INVENTION: CONSTITUTIVE EXPRESSION OF P450SOY
TITLE OF INVENTION: EXPRESSION OF P450SOY
TITLE OF INVENTION: AND FERREDOXIN-SOY IN
TITLE OF INVENTION: STREPTOMYCES
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: E. I. DU PONT DE NEMOURS
ADDRESSEE: AND COMPANY
STREET: 1007 MARKET STREET
CITY: WILMINGTON
STATE: DELAWARE
COUNTRY: USA
ZIP: 19898
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch,
MEDIUM TYPE: 1.0 MB
```

COMPUTER: Macintosh  
OPERATING SYSTEM: Macintosh System, 6.0  
SOFTWARE: Microsoft Word, 4.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US92/10885  
FILING DATE: 19921226  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: GALLAGOS, R. THOMAS  
REGISTRATION NUMBER: 32,692  
REFERENCE/DOCKET NUMBER: CR-9000-A  
TELEPHONE: 302-892-7342  
TELEFAX: 302-892-7949  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 412 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
PCT-US92-10885-11

Query Match 14.5%; Score 315; DB 5; Length 412;  
Best Local Similarity 27.2%; Pred. No. 3e-24;  
Matches 84; Conservative 56; Mismatches 163; Indels 6; Gaps 5;

QY 103 MDPPEQRFALANQVGMVVDKLENRIQELACSLIESLRQCG-QCNFTEDYAEPPPIR 161  
Db 105 VDFEHNTRQRMILPTFSVKRIGALRPRIQVTDRLDAMERQGPFAELVSAFALVPSM 164  
QY 162 IFMLLAGLFEEDI PHLKYLTDQMTDPGSMTFABAKAALDYLIPIIQRQKPKQTDALS 221  
Db 165 VICALLGVYADHAFFERSQRLLEGPCADQVNRARDEBEYLGLIDRKRAEPDGLLD 224  
QY 222 IVANGQVNGRPIITDEAKMCGLLLVGLDITVWVFLSFMFLAKSPHROELIERPERI 281  
Db 225 ELIRDHEDGVDRQCLVAFVILLIAGHETTANNISLGTFTLLSHPEQLAALRAGGTST 284  
QY 282 PAACEELLRRFSVADG--RLTSDYEFHGVQLKGGDQILPQLMSGLDERENACPMHVD 339  
Db 285 AVVTEELL-RLSLAEGQLRLATDMEVDGATIRKGGVWFSTSLINRDADVFPRAETLD 343  
QY 340 FSRQKVSHTTFHGHSLCLGQLHARREIIVTLKEWLTTRIPDFSIA-PGAQI-CHKSG-IVS 397  
Db 344 WDRPARHLAFGFGVHQLGQNLARAEIDIAMRTLFEPLGRLAVPAHEIRHKGDTIQ 403  
QY 398 GVQALPLVW 406  
Db 404 GLLDLPVAV 412

RESULT 14  
US-09-335-409-8  
; Sequence 8, Application US/09335409  
; Patent No. 6121029  
; GENERAL INFORMATION:  
; APPLICANT: Schupp, Thomas  
; APPLICANT: Ligon, James  
; APPLICANT: Molnar, Istvan  
; APPLICANT: Zirkle, Ross  
; APPLICANT: Cyr, Devon  
; APPLICANT: Goerlach, Joern  
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES  
; FILE REFERENCE: 4-30582A  
; CURRENT APPLICATION NUMBER: US/09/335,409  
; CURRENT FILING DATE: 1999-06-17  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 8  
; LENGTH: 419  
; TYPE: PRT  
; ORGANISM: Sorangium cellulosum

US-09-335-409-8

Query Match 12.5%; Score 271.5; DB 3; Length 419;  
Best Local Similarity 23.8%; Pred. No. 1.4e-19;  
Matches 100; Conservative 67; Mismatches 179; Indels 75; Gaps 14;

QY 6 QSNANLAPPHVPEHLVFFDFMYPNLSAGVQE-----AWAVLQESNVDPDLVWTRCNGG 61  
Db 3 QEQANQSETKP-----AFDFKFPAP-----GVAEDPFPALERLREA-TPIFYD--EGR 48  
QY 62 HMLATSEQ-----LIREAYEDYRHFSSBQFPFIPREAGEAYDFIPTSMDDPPEQR 109  
Db 49 SWLTRYHDVSAVERDERFAVSREWESEAESSAIZ-----ELSDMKKYGLFGLPPSDHA 104  
QY 110 QFRALANQVGMVVDKLENRIQELACSLIESLRQCGQCNFTEDYAEPPPIRIFMLLAGL 169  
Db 105 RYRKLVNPSFTSRAIDLLRAEIQRTVDQLDARSQGEFFDVVDYAEGIPMRAISALLKV 164  
QY 170 PREDIPHLKYLTDQMTDPGSMTFABAKAALDYLIPI-----I 207  
Db 165 PAE-----CDEKFRFRFSAT-----ARALGVGLVQVQVDEETKTLVASVTEGLALLHDV 212  
QY 208 IFORQKQ2-GTDAISIVANGQVNGRPIITSDAKRNCGLLVGLDITVWVFLSFMFLAK 266  
Db 213 LDERRNPLENDVLTMLLQAEADGSRSLSTKELVALVGAIIAAGTETTYLIIFAVLNLLR 272  
QY 267 SPEHRQELIERPERIPAAACEELLRRFSVADG--RLTSDYEFHGVQLKGGDQI---LLPQ 322  
Db 273 SPEALELVKAEPLMERNALDEVLRFDNLRITVTRFARQDLEYCGASIKKGEVFFLLIFS 332  
QY 323 MSLGLDERENACPMHVDPSRQKVSHTTFHGHSLCLGQLHARREIIVTLKEWLTTRIPDFS 382  
Db 333 ALR--DGTVFSREDVDFVDRDGTASLAYGRGPHVCGVSLARLEAEIAGTIFRRPEMK 390  
QY 383 I 383  
Db 391 L 391

RESULT 15  
US-09-413-814-71  
; Sequence 71, Application US/09413814  
; Patent No. 6225664  
; GENERAL INFORMATION:  
; APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH  
; APPLICANT: Bristol-Myers Squibb, Co.  
; APPLICANT: Beyer, Stefan  
; APPLICANT: Bloeker, Helmut  
; APPLICANT: Brandt, Petra  
; APPLICANT: Cino, Paul M  
; APPLICANT: Dougherty, Brian A  
; APPLICANT: Goldberg, Steven L  
; APPLICANT: Hefle, Gerhard  
; APPLICANT: Mueller, Joachim  
; APPLICANT: Reichenbach, Hans  
; TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polyketide or  
; TITLE OF INVENTION: heteropolyketide compounds  
; FILE REFERENCE: PCT/US 99/23535  
; CURRENT APPLICATION NUMBER: US/09/413,814  
; CURRENT FILING DATE: 1999-10-07  
; EARLIER APPLICATION NUMBER: DE 198 46 493.2  
; EARLIER FILING DATE: 1998-10-09  
; NUMBER OF SEQ ID NOS: 107  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 71  
; LENGTH: 419  
; TYPE: PRT  
; ORGANISM: Sorangium cellulosum  
US-09-413-814-71

Query Match 12.5%; Score 271.5; DB 3; Length 419;  
Best Local Similarity 23.8%; Pred. No. 1.4e-19;  
Matches 100; Conservative 67; Mismatches 179; Indels 75; Gaps 14;





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 6, 2004, 18:53:57 ; Search time 14.974 Seconds

(without alignments)  
3108.883 Million cell updates/sec

Title: US-09-246-451A-11

Perfect score: 2180

Sequence: 1 ITETIQSNANLAPPHVPE.....IVSGVQALPLYWDPAITTKAV 414

Scoring table: BLOSUM62

Gapop 13.0 , Gapext 0.5

Searched: 1071772 seqs, 262633353 residues

Total number of hits satisfying chosen parameters: 1071772

Minimum DB seq length: 3

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.psp.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.psp.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.psp.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.psp.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.psp.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.psp.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.psp.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.psp.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.psp.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.psp.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.psp.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.psp.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.psp.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.psp.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.psp.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10C\_NEW\_PUB.psp.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.psp.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.psp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2180	100.0	414	15	US-10-453-104-11
2	2176	99.8	414	15	US-10-453-104-11
3	2173	99.7	414	15	US-10-453-104-12
4	2169	99.5	414	15	US-10-453-104-13
5	377.5	17.3	404	14	US-10-214-446-50
6	376.5	17.3	404	14	US-10-214-446-40
7	372	17.1	399	14	US-10-314-657-37
8	367	16.8	416	14	US-10-156-761-14997
9	365.5	16.8	416	9	US-09-861-289-39
10	365.5	16.8	416	9	US-09-860-846-39
11	365.5	16.8	416	10	US-09-988-384B-39
12	365.5	16.8	416	10	US-09-836-821-39
13	365.5	16.8	416	10	US-09-793-708-18
14	365.5	16.8	416	14	US-10-201-365-13
15	365.5	16.8	416	14	US-10-160-539-18

16	365.5	16.8	416	14	US-10-271-889-39	Sequence 39, Appl
17	363.5	16.7	411	14	US-10-156-761-8376	Sequence 8376, Ap
18	358	16.4	393	14	US-10-156-761-9525	Sequence 9525, Ap
19	357	16.4	405	14	US-10-214-446-38	Sequence 38, Appl
20	356	16.3	425	14	US-10-214-446-20	Sequence 20, Appl
21	354.5	16.3	457	14	US-10-156-761-11073	Sequence 11073, A
22	351.5	16.2	399	14	US-10-156-761-9914	Sequence 9914, Ap
23	350	16.1	418	12	US-10-389-647-559	Sequence 559, Ap
24	341.5	15.7	388	14	US-10-156-761-13776	Sequence 13776, A
25	338.5	15.5	408	14	US-10-214-446-4	Sequence 4, Appli
26	333	15.3	409	15	US-10-458-201-12	Sequence 12, Appl
27	332	15.2	392	14	US-10-214-446-32	Sequence 32, Appl
28	328.5	15.1	406	14	US-10-214-446-2	Sequence 2, Appli
29	322	14.8	401	14	US-10-156-761-8710	Sequence 8710, Ap
30	321	14.7	404	14	US-10-214-446-16	Sequence 16, Appl
31	318.5	14.6	396	14	US-10-214-446-14	Sequence 14, Appl
32	318	14.6	428	14	US-10-201-213-6	Sequence 6, Appli
33	317.5	14.6	404	14	US-10-156-761-14659	Sequence 14659, A
34	315	14.4	412	14	US-10-214-446-36	Sequence 36, Appl
35	315	14.4	421	14	US-10-156-761-9703	Sequence 9703, Ap
36	314	14.4	418	14	US-10-214-446-22	Sequence 22, Appl
37	307	14.1	421	14	US-10-214-446-18	Sequence 18, Appl
38	304.5	14.0	404	14	US-10-156-761-10431	Sequence 10431, A
39	302.5	13.9	415	14	US-10-214-446-46	Sequence 46, Appl
40	301.5	13.8	400	14	US-10-314-657-62	Sequence 62, Appl
41	298.5	13.7	402	14	US-10-205-032-8	Sequence 8, Appli
42	298	13.7	475	14	US-10-145-415-22	Sequence 22, Appl
43	297	13.6	470	14	US-10-145-415-6	Sequence 6, Appli
44	295.5	13.6	430	9	US-09-738-626-4117	Sequence 4117, Ap
45	295	13.5	399	14	US-10-156-761-7959	Sequence 7959, Ap

ALIGNMENTS

RESULT 1

US-10-453-104-11  
; Sequence 11, Application US/10453:04  
; Publication No. US20030207345A1  
; GENERAL INFORMATION:  
; APPLICANT: California Institute of Technology;  
; APPLICANT: Frances H. Arnold  
; APPLICANT: Hyun Joo  
; TITLE OF INVENTION: Oxygenase Enzymes and Screening Method  
; FILE REFERENCE: 4058/1E827-US3  
; CURRENT APPLICATION NUMBER: US/10/453.104  
; PRIOR FILING DATE: 2003-06-02  
; PRIOR APPLICATION NUMBER: US 09/661,093  
; PRIOR FILING DATE: 2000-09-13  
; PRIOR APPLICATION NUMBER: US 09/246,451  
; PRIOR FILING DATE: 1999-02-09  
; PRIOR APPLICATION NUMBER: US 60/094,403  
; PRIOR FILING DATE: 1998-07-28  
; PRIOR APPLICATION NUMBER: US 60/106,840  
; PRIOR FILING DATE: 1998-11-03  
; PRIOR APPLICATION NUMBER: US 60/086,206  
; PRIOR FILING DATE: 1998-05-21  
; PRIOR APPLICATION NUMBER: US 60/106,834  
; PRIOR FILING DATE: 1998-11-03  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 11  
; LENGTH: 414  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Mutant M7-4H  
US-10-453-104-11

Query Match 100.0%; Score 2180; DB 15; Length 414;  
Best Local Similarity 100.0%; Pred. No. 7,1e-212;  
Matches 414; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 TTETIQSNANLAPLPHVPEHLVDFDMYNPSNLSAGVQEAVALQESNVPDLVWTRCNG 60
DB 1 TTETIQSNANLAPLPHVPEHLVDFDMYNPSNLSAGVQEAVALQESNVPDLVWTRCNG 60
QY 61 GHWIATRGQIIRAEYEDYRHFSSCEPPIPREAGEAYDFIPTSMDFPQORFRALANQVVG 120
DB 61 GHWIATRGQIIRAEYEDYRHFSSCEPPIPREAGEAYDFIPTSMDFPQORFRALANQVVG 120
QY 121 MPVVDKLENRIQELACSLIESLRPQGCNFTEDYAEPPPIRIFMLLAGLPEEDIPHLKYL 180
DB 121 MPVVDKLENRIQELACSLIESLRPQGCNFTEDYAEPPPIRIFMLLAGLPEEDIPHLKYL 180
QY 181 TDQWTRPDGSMTEFAEAKAALYDYLPIIEQRORKEGTDAISIVANGQVNGRPITSDAKR 240
DB 181 TDQWTRPDGSMTEFAEAKAALYDYLPIIEQRORKEGTDAISIVANGQVNGRPITSDAKR 240
QY 241 MCGLLLVGGGLDVTNNFLSFMSEFLAKSPEHRQELIERPERIPAAACELLRRFSLVADGRI 300
DB 241 MCGLLLVGGGLDVTNNFLSFMSEFLAKSPEHRQELIERPERIPAAACELLRRFSLVADGRI 300
QY 301 LTSDFEFGVQLKGGDQILLPOMLSGLDERKNACPMHVDFSRQKVSHTTFEGSHLCLGQ 360
DB 301 LTSDFEFGVQLKGGDQILLPOMLSGLDERKNACPMHVDFSRQKVSHTTFEGSHLCLGQ 360
QY 361 HLARREIIVTLKEWLTRIPDFSAPGAQIOHKSGIVSGVQALPLVWDPATTKAV 414
DB 361 HLARREIIVTLKEWLTRIPDFSAPGAQIOHKSGIVSGVQALPLVWDPATTKAV 414

```

## RESULT 2

```

US-10-453-104-2
; Sequence 2, Application US/10453104
; Publication No. US20030207345A1
; GENERAL INFORMATION:
; APPLICANT: California Institute of Technology;
; APPLICANT: Frances H. Arnold
; APPLICANT: Hyun Joo
; TITLE OF INVENTION: Oxygenase Enzymes and Screening Method
; FILE REFERENCE: 4058/1E827-US3
; CURRENT APPLICATION NUMBER: US 10/453,104
; CURRENT FILING DATE: 2003-06-02
; PRIOR APPLICATION NUMBER: US 09/661,093
; PRIOR FILING DATE: 2000-09-13
; PRIOR APPLICATION NUMBER: US 05/246,451
; PRIOR FILING DATE: 1998-02-29
; PRIOR APPLICATION NUMBER: US 60/094,403
; PRIOR FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: US 60/106,840
; PRIOR FILING DATE: 1998-11-03
; PRIOR APPLICATION NUMBER: US 60/086,206
; PRIOR FILING DATE: 1998-05-21
; PRIOR APPLICATION NUMBER: US 60/106,834
; PRIOR FILING DATE: 1998-11-03
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Fast-Seq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 414
; TYPE: PRT
; ORGANISM: P. Putida
US-10-453-104-2

```

```

Query Match 99.8%; Score 2176; DB 15; Length 414;
Best Local Similarity 99.8%; Pred. No. 1.8e-211;
Matches 413; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTETIQSNANLAPLPHVPEHLVDFDMYNPSNLSAGVQEAVALQESNVPDLVWTRCNG 60
DB 1 TTETIQSNANLAPLPHVPEHLVDFDMYNPSNLSAGVQEAVALQESNVPDLVWTRCNG 60
QY 61 GHWIATRGQIIRAEYEDYRHFSSCEPPIPREAGEAYDFIPTSMDFPQORFRALANQVVG 120
DB 61 GHWIATRGQIIRAEYEDYRHFSSCEPPIPREAGEAYDFIPTSMDFPQORFRALANQVVG 120

```

```

QY 121 MPVVDKLENRIQELACSLIESLRPQGCNFTEDYAEPPPIRIFMLLAGLPEEDIPHLKYL 180
DB 121 MPVVDKLENRIQELACSLIESLRPQGCNFTEDYAEPPPIRIFMLLAGLPEEDIPHLKYL 180
QY 181 TDQWTRPDGSMTEFAEAKAALYDYLPIIEQRORKEGTDAISIVANGQVNGRPITSDAKR 240
DB 181 TDQWTRPDGSMTEFAEAKAALYDYLPIIEQRORKEGTDAISIVANGQVNGRPITSDAKR 240
QY 241 MCGLLLVGGGLDVTNNFLSFMSEFLAKSPEHRQELIERPERIPAAACELLRRFSLVADGRI 300
DB 241 MCGLLLVGGGLDVTNNFLSFMSEFLAKSPEHRQELIERPERIPAAACELLRRFSLVADGRI 300
QY 301 LTSDFEFGVQLKGGDQILLPOMLSGLDERKNACPMHVDFSRQKVSHTTFEGSHLCLGQ 360
DB 301 LTSDFEFGVQLKGGDQILLPOMLSGLDERKNACPMHVDFSRQKVSHTTFEGSHLCLGQ 360
QY 361 HLARREIIVTLKEWLTRIPDFSAPGAQIOHKSGIVSGVQALPLVWDPATTKAV 414
DB 361 HLARREIIVTLKEWLTRIPDFSAPGAQIOHKSGIVSGVQALPLVWDPATTKAV 414

```

## RESULT 3

```

US-10-453-104-12
; Sequence 12, Application US/10453104
; Publication No. US20030207345A1
; GENERAL INFORMATION:
; APPLICANT: California Institute of Technology;
; APPLICANT: Frances H. Arnold
; APPLICANT: Hyun Joo
; TITLE OF INVENTION: Oxygenase Enzymes and Screening Method
; FILE REFERENCE: 4058/1E827-US3
; CURRENT APPLICATION NUMBER: US 10/453,104
; CURRENT FILING DATE: 2003-06-02
; PRIOR APPLICATION NUMBER: US 09/661,093
; PRIOR FILING DATE: 2000-09-13
; PRIOR APPLICATION NUMBER: US 09/246,451
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: US 60/094,403
; PRIOR FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: US 60/106,840
; PRIOR FILING DATE: 1998-11-03
; PRIOR APPLICATION NUMBER: US 60/086,206
; PRIOR FILING DATE: 1998-05-21
; PRIOR APPLICATION NUMBER: US 60/106,834
; PRIOR FILING DATE: 1998-11-03
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mutant M7-6H
US-10-453-104-12

```

```

Query Match 99.7%; Score 2173; DB 15; Length 414;
Best Local Similarity 99.8%; Pred. No. 3.6e-211;
Matches 413; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 TTETIQSNANLAPLPHVPEHLVDFDMYNPSNLSAGVQEAVALQESNVPDLVWTRCNG 60
DB 1 TTETIQSNANLAPLPHVPEHLVDFDMYNPSNLSAGVQEAVALQESNVPDLVWTRCNG 60
QY 61 GHWIATRGQIIRAEYEDYRHFSSCEPPIPREAGEAYDFIPTSMDFPQORFRALANQVVG 120
DB 61 GHWIATRGQIIRAEYEDYRHFSSCEPPIPREAGEAYDFIPTSMDFPQORFRALANQVVG 120
QY 121 MPVVDKLENRIQELACSLIESLRPQGCNFTEDYAEPPPIRIFMLLAGLPEEDIPHLKYL 180
DB 121 MPVVDKLENRIQELACSLIESLRPQGCNFTEDYAEPPPIRIFMLLAGLPEEDIPHLKYL 180
QY 181 TDQWTRPDGSMTEFAEAKAALYDYLPIIEQRORKEGTDAISIVANGQVNGRPITSDAKR 240

```

```
Db 181 TDQMTREDGSMFTAENKALYDYLPIIEQRQKPGTDAISIVNGQVNGRPITSDEAKR 240
Qy 241 MCGLLLVGGLETVVNFSLFSMEFLAKSPHQRQELIERPERIPAACEBELLRRFSLVADGRI 300
Db 241 MCGLLLVGGLETVVNFSLFSMEFLAKSPHQRQELIERPERIPAACEBELLRRFSLVADGRI 300
Qy 301 LTSDFEFGVQLKKGQDILLPQMLSGLDERKNACPMHVDPSRQKVSHTTFGSHGLCLGQ 360
Db 301 LTSDFEFGVQLKKGQDILLPQMLSGLDERKNACPMHVDPSRQKVSHTTFGSHGLCLGQ 360
Qy 361 HJARREIIVTKEWLIRIPDFSAPGAQIOHKSGIVSGVQALPLVMDPATTKAV 414
Db 361 HJARREIIVTKEWLIRIPDFSAPGAQIOHKSGIVSGVQALPLVMDPATTKAV 414

RESULT 4
US-10-453-104-13
; Sequence 13, Application US/10453104
; Publication No. US20030207345A1
; GENERAL INFORMATION:
; APPLICANT: California Institute of Technology;
; APPLICANT: Frances H. Arnold
; APPLICANT: Hyun Joo
; TITLE OF INVENTION: Oxxygenase Enzymes and Screening Method
; FILE REFERENCE: 4058/1827-US3
; CURRENT APPLICATION NUMBER: US/10/453,104
; CURRENT FILING DATE: 2003-06-02
; PRIOR APPLICATION NUMBER: US 09/661,093
; PRIOR FILING DATE: 2000-09-13
; PRIOR APPLICATION NUMBER: US 09/246,451
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: US 60/394,403
; PRIOR FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: US 60/106,840
; PRIOR FILING DATE: 1998-11-03
; PRIOR APPLICATION NUMBER: US 60/686,206
; PRIOR FILING DATE: 1998-05-21
; PRIOR APPLICATION NUMBER: US 60/106,834
; PRIOR FILING DATE: 1998-11-03
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mutant M7-8H
US-10-453-104-13

Query Match 99.5%; Score 2169; DB 15; Length 414;
Best Local Similarity 99.8%; Pred. No. 9.3e-211;
Matches 413; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TTETIQSANLAPPPHVEHLVDFDVMYNLSNLSAGVQEAVALQESNVPLVWTRCNG 60
Db 1 TTETIQSANLAPPPHVEHLVDFDVMYNLSNLSAGVQEAVALQESNVPLVWTRCNG 60
Qy 61 GHVIATRGQIREAYEDYRHSSECFPIPRAGEAYDFTPTSDPPQORQFALANQVVG 120
Db 61 GHVIATRGQIREAYEDYRHSSECFPIPRAGEAYDFTPTSDPPQORQFALANQVVG 120
Qy 121 MPVVDKLENRQELACSLIESLRPOGNETEDYAEPPPIRIFMLLAGLPEEDIPLHKYL 180
Db 121 MPVVDKLENRQELACSLIESLRPOGNETEDYAEPPPIRIFMLLAGLPEEDIPLHKYL 180
Qy 181 TDQMTREPGSMFTAENKALYDYLPIIEQRQKPGTDAISIVANGQVNGRPITSDEAKR 240
Db 181 TDQMTREPGSMFTAENKALYDYLPIIEQRQKPGTDAISIVANGQVNGRPITSDEAKR 240
Qy 241 MCGLLLVGGLETVVNFSLFSMEFLAKSPHQRQELIERPERIPAACEBELLRRFSLVADGRI 300
Db 241 MCGLLLVGGLETVVNFSLFSMEFLAKSPHQRQELIERPERIPAACEBELLRRFSLVADGRI 300
```

```
Qy 301 LTSDFEFGVQLKKGQDILLPQMLSGLDERKNACPMHVDPSRQKVSHTTFGSHGLCLGQ 360
Db 301 LTSDFEFGVQLKKGQDILLPQMLSGLDERKNACPMHVDPSRQKVSHTTFGSHGLCLGQ 360
Qy 361 HJARREIIVTKEWLIRIPDFSAPGAQIOHKSGIVSGVQALPLVMDPATTKAV 414
Db 361 HJARREIIVTKEWLIRIPDFSAPGAQIOHKSGIVSGVQALPLVMDPATTKAV 414

RESULT 5
US-10-214-446-50
; Sequence 50, Application US/10214446
; Publication No. US20030180742A1
; GENERAL INFORMATION:
; APPLICANT: Weiner, David
; APPLICANT: Burk, Mark J.
; APPLICANT: Hitchman, Tim
; APPLICANT: Fujol, Catherine
; APPLICANT: Richardsort, Toby
; APPLICANT: Short, Jay M.
; TITLE OF INVENTION: P450 ENZYMES, NUCLEIC ACIDS ENCODING
; TITLE OF INVENTION: THEM AND METHODS OF MAKING AND USING THEM
; FILE REFERENCE: 09010-5003C1
; CURRENT APPLICATION NUMBER: US/10/214,446
; CURRENT FILING DATE: 2002-08-05
; PRIOR APPLICATION NUMBER: US 60/309,497
; PRIOR FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 50
; LENGTH: 404
; TYPE: PRT
; ORGANISM: Bacterial
US-10-214-446-50

Query Match 17.3%; Score 377.5; DB 14; Length 404;
Best Local Similarity 27.6%; Pred. No. 3.6e-29;
Matches 97; Conservative 64; Mismatches 157; Indels 33; Gaps 5;

Qy 60 GGHWTATRGQIREAYEDYRHSSECFPIPRAGEAYDFTPTSM-----103
Db 32 GRTWFLPHHADIRTAIRDRFSAS-----RYAGGVNQFFAEVPEFAFNEAISRWIV 85
Qy 104 --DPPEQRQFALANQVVGMPVVDKLENRIQELACSLIESLRPOGNETEDYAEPPPIR 161
Db 86 LHDQPEHQLRQLMQGGFTRLITTMPEKIORVCDLIDAFVKRGSTEFMTYAHPEPAK 145
Qy 162 IFMLLAGLPEEDIPLHKYLTDQMTREPGSM-----TFEAKE---ALYDYLPIIEQRQ 213
Db 146 VIAEMGVNPDYPAPFVVMSEDLNFAFAGSLRPTLEMFRAAQDGLLNMDFARLLPERE 205
Qy 214 KPGTDAISIVANGQVNGRPITSDEAKRMCGLLLVGGLETVVNFSLFSMEFLAKSPHQRQ 273
Db 206 NEGDDLVSLLLSAEGEWMTAEQVLANTQIIVAGHETTRNLVANGVELLRYPEQAL 265
Qy 274 LTERPERIPAACEBELLRRFS-LVADGRILTSYEHFGVQLKKGQDILLQMLSGLDERKN 332
Db 266 LESRELMPSAVREIMRFESPLQIFRRVAREDFEGGAEREGDGLVLMGSAANDPEAF 325
Qy 333 ACPMHVDPSRQKVSHTTFGSHGLCLGQHLARREIIVTKEWLIRIPDFS 383
Db 326 DDPTDFDLTRNPTGTAFGWGPHVCVGAALAELEGQVSPFRTLLDRPLGLE 376

RESULT 6
US-10-214-446-40
; Sequence 40, Application US/10214446
; Publication No. US20030180742A1
; GENERAL INFORMATION:
; APPLICANT: Weiner, David
; APPLICANT: Burk, Mark J.
; APPLICANT: Hitchman, Tim
; APPLICANT: Fujol, Catherine
```

```

US-10-314-659-37

Query Match      17.1%; Score 372; DB 14; Length 399;
Best Local Similarity 28.1%; Pred. No. 1.3e-28;
Matches 1.0; Conservative 70; Mismatches 176; Indels 36; Gaps 9;

QY 38 VOEWAVLQESNVVDLVWTRCNGHMIATRGQILREAYEDYRHFSBCCPIPREAGE--- 94
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 21 IHPKFAELRETDPLARVRLPYGCGEWMVTR-----YDDVRAANSDFRSRAQIGDTP 73
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 95 -----AYDFI2SDMPPEQORFALANQVGMPPVVDKLENIQELACSLIIESLRPQGQ 147
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 74 RTTPELARSDTI-LSLDPPPEHTERLLSKAFTARRMGAMQSWLEELFAGLLDGVETGH 132
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 148 -CNFTEDYAPFPPIRIFMLIAGLPEEDIPLHKYLTIDQMTRPDGSMTFAEAKE-----A 199
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 133 PADIVRDLAQPFITAVICLLGWPYEDRGFQHWSEVI-----MSTTAYSKEEAASADAS 187
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 200 LYDYLPIITIBQRQKPGTDAISVANGQVNGRPITSDAEKMCLLIAGGLDVTWNFLSF 259
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 188 IRAYLADLVSARAAAPHDDLLGLVVSARDDDRLTEDELITFGVTLLVAGHETSAHQGN 247
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 260 SMEFLAKSPHRQELTERPERIPAAACEILARRSL---VADGRILTSDYEPHGVLQAKGD 316
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 248 MVYALLTHEDOLSLRQPELLPRABEELRFPVLGNGVGNARIALEDVLSGGTVRAGE 307
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 317 QILLPOMLSGLDERKNACPMHEVPFSRQKVSHTTTFGHGSHLCLGOHLARREIIVTLKEMLT 376
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 308 GVVAAVANNRDPRAFDDPRLDITREKNPHLAFHGHAHYCLGAQLARMELRVAIGELLE 367
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 377 RIDDFSTA-PCAQIQHKS-GIVSGVQALPLVW 406
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 368 RFGGLRLAVPADQVQEWKTTGGLFGRQELPTAW 399
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15103
; SEQ ID NO 14997
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-14997

Query Match      16.8%; Score 367; DB 14; Length 416;
Best Local Similarity 29.1%; Pred. NO. 4.4e-28;
Matches 104; Conservative 56; Mismatches 181; Indels 16; Gaps 8

QY      63  WIAIRGQLIREAYEDYRHFSCECPPIR--EAGEAYDEIPT-----SMDPEQQRFRAL 114
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      63  WVVTGHAARALLSDRLSSDRT--LPFPATTEFEAVRTFRVALLGVDDPEHRTQRM 120

QY      115  ANQVVGMFVVDKLENRIQELACSLIESLRQG--QNFTEDYAEPPPIRIFMLLAGLPEED 173
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      121  LVPSEFTLKRAAALRRIQETVDGLIDAMEAQGFPAELVSAPFLPESVMVICALLGVPYAL 180

```

RESULT 10  
US-09-860-346-39  
: Sequence 39. Application US/09860846

```

RESULT 11
US-09-988-384B-39
Sequence 39, Application US/09988384B
Publication NO. US20030073824A1
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
APPLICANT: Zhao, L.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600.536US1
CURRENT APPLICATION NUMBER: US/09/988,384B
CURRENT FILING DATE: 2001-11-19
PRIOR APPLICATION NUMBER: PCT/US99/14398
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: US 09/105,537
PRIOR FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 53
SEQ ID NO 39
LENGTH: 416
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-988-384B-39

```

Query Match 16.8%; Score 365.5; DB 10; Length 416;  
Best Local Similarity 28.3%; Pred. No. 6.2e-28;  
Matches 97; Conservative 62; Mismatches 163; Indels 21; Gaps 7;

QY 76 EDYRHSFSECPPIPREAGEAYDFTPTSDPPEQRFALANQVGMVVDKLENRIQELA 135  
DB 72 KQWR--NSTTPTLEAEALNNHML--DPPHTRKRLVAREFTMRVVELLRPRVQEV 127

QY 136 CSLIESL--RPQGCNFTEDYAEPPPIRIFMLLAGLPEEDIPHLKYLTDQMTDPGSMTF 193  
DB 128 DGLVDMLAAPDGRADLMESLAWPLITVISSELLGVPEPDRAAFRVWIDAFVFPDPAQA 187

QY 194 ABAEALYDYLIPIEQRQKPGTDAIS--IVANGVNGRPITSDKAKMCGLLVGGT 252  
DB 188 QTAMAEMSGYLSRLIDSKRGQDGEDLLSALVTSDEDSRLTSEELLGMAHILLVAGHET 247

QY 253 VVNFLSFMEFLAKSPEHRQELIERPERIPACEELLRRFSLVADGRILTSDFE----- 307  
DB 248 TVNLITANGYALLSHPDQALRADMTLLDGAVEMLR-----YEGPVESATYRFPV 302

QY 308 --HGVLKKGQDQILLPQMLSGLDERKNACPMHVDFSRQKVSHTTTFHGSHLCLGQHLARR 365  
DB 303 DLDGTVPAGDTLVVLADAHRTPERFDPHRRDIRDTAGHLAFGHGHCIGAPLARR 362

QY 366 EIIVTLKEMLTRIPDFS--IAPGAQIQHKSGLVSGVQALPLVW 406  
DB 363 EARIAVALLERCPLDALDVSPGELVWYPMIRGLKALPIRW 405

RESULT 13  
US-09-793-708-18  
; Sequence 18, Application US/09793708  
; Publication No. US20030104597A1  
; GENERAL INFORMATION:  
; APPLICANT: ASHLEY, Gary  
; APPLICANT: BETLACH, Melanie C.  
; APPLICANT: BETLACH, Mary C.  
; APPLICANT: MCDANIEL, Robert  
; APPLICANT: TANG, Li  
; TITLE OF INVENTION: RECOMBINANT NARBONCLIDE POLYKETIDE SYNTHASE  
; FILE REFERENCE: 300622002121  
; CURRENT APPLICATION NUMBER: US/09/793,708  
; CURRENT FILING DATE: 2001-02-22  
; PRIOR APPLICATION NUMBER: US 09/657,440  
; PRIOR FILING DATE: 2000-09-07  
; PRIOR APPLICATION NUMBER: US 09/320,878  
; PRIOR FILING DATE: 1999-05-27  
; PRIOR APPLICATION NUMBER: US 09/141,908  
; PRIOR FILING DATE: 1998-08-28  
; PRIOR APPLICATION NUMBER: US 09/073,538  
; PRIOR FILING DATE: 1998-05-06  
; PRIOR APPLICATION NUMBER: US 08/846,247  
; PRIOR FILING DATE: 1997-04-30  
; PRIOR APPLICATION NUMBER: US 60/134,990  
; PRIOR FILING DATE: 1999-05-20  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: Patent in ver. 2.0  
; SEQ ID NO 18  
; LENGTH: 416  
; TYPE: PRT  
; ORGANISM: Streptomyces venezuelae  
US-09-793-708-18

Query Match 16.8%; Score 365.5; DB 10; Length 416;  
Best Local Similarity 28.3%; Pred. No. 6.2e-28;  
Matches 97; Conservative 62; Mismatches 163; Indels 21; Gaps 7;

QY 76 EDYRHSFSECPPIPREAGEAYDFTPTSDPPEQRFALANQVGMVVDKLENRIQELA 135  
DB 72 KQWR--NSTTPTLEAEALNNHML--DPPHTRKRLVAREFTMRVVELLRPRVQEV 127

QY 136 CSLIESL--RPQGCNFTEDYAEPPPIRIFMLLAGLPEEDIPHLKYLTDQMTDPGSMTF 193  
DB 128 DGLVDMLAAPDGRADLMESLAWPLITVISSELLGVPEPDRAAFRVWIDAFVFPDPAQA 187

QY 194 ABAEALYDYLIPIEQRQKPGTDAIS--IVANGVNGRPITSDKAKMCGLLVGGT 252  
DB 188 QTAMAEMSGYLSRLIDSKRGQDGEDLLSALVTSDEDSRLTSEELLGMAHILLVAGHET 247

QY 253 VVNFLSFMEFLAKSPEHRQELIERPERIPACEELLRRFSLVADGRILTSDFE----- 307  
DB 248 TVNLITANGYALLSHPDQALRADMTLLDGAVEMLR-----YEGPVESATYRFPV 302

QY 308 --HGVLKKGQDQILLPQMLSGLDERKNACPMHVDFSRQKVSHTTTFHGSHLCLGQHLARR 365  
DB 303 DLDGTVPAGDTLVVLADAHRTPERFDPHRRDIRDTAGHLAFGHGHCIGAPLARR 362

QY 366 EIIVTLKEMLTRIPDFS--IAPGAQIQHKSGLVSGVQALPLVW 406  
DB 363 EARIAVALLERCPLDALDVSPGELVWYPMIRGLKALPIRW 405

RESULT 14  
US-10-201-365-13

Query Match 16.8%; Score 365.5; DB 10; Length 416;  
Best Local Similarity 28.3%; Pred. No. 6.2e-28;  
Matches 97; Conservative 62; Mismatches 163; Indels 21; Gaps 7;

QY 76 EDYRHSFSECPPIPREAGEAYDFTPTSDPPEQRFALANQVGMVVDKLENRIQELA 135  
DB 72 KQWR--NSTTPTLEAEALNNHML--DPPHTRKRLVAREFTMRVVELLRPRVQEV 127

QY 136 CSLIESL--RPQGCNFTEDYAEPPPIRIFMLLAGLPEEDIPHLKYLTDQMTDPGSMTF 193  
DB 128 DGLVDMLAAPDGRADLMESLAWPLITVISSELLGVPEPDRAAFRVWIDAFVFPDPAQA 187

QY 194 ABAEALYDYLIPIEQRQKPGTDAIS--IVANGVNGRPITSDKAKMCGLLVGGT 252  
DB 188 QTAMAEMSGYLSRLIDSKRGQDGEDLLSALVTSDEDSRLTSEELLGMAHILLVAGHET 247

QY 253 VVNFLSFMEFLAKSPEHRQELIERPERIPACEELLRRFSLVADGRILTSDFE----- 307  
DB 248 TVNLITANGYALLSHPDQALRADMTLLDGAVEMLR-----YEGPVESATYRFPV 302

QY 308 --HGVLKKGQDQILLPQMLSGLDERKNACPMHVDFSRQKVSHTTTFHGSHLCLGQHLARR 365  
DB 303 DLDGTVPAGDTLVVLADAHRTPERFDPHRRDIRDTAGHLAFGHGHCIGAPLARR 362

QY 366 EIIVTLKEMLTRIPDFS--IAPGAQIQHKSGLVSGVQALPLVW 406  
DB 363 EARIAVALLERCPLDALDVSPGELVWYPMIRGLKALPIRW 405

RESULT 12  
US-09-836-821-39  
; Sequence 39, Application US/09836821  
; Publication No. US20030037405A1  
; GENERAL INFORMATION:  
; APPLICANT: Shermat, D. H.  
; APPLICANT: Liu, H.  
; APPLICANT: Xue, Y.  
; APPLICANT: Zhao, Z.  
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin  
; FILE REFERENCE: 600.43801  
; CURRENT APPLICATION NUMBER: US/09/836,821  
; CURRENT FILING DATE: 2001-04-17  
; PRIOR APPLICATION NUMBER: 09/105,537  
; PRIOR FILING DATE: 1998-06-26  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 39  
; LENGTH: 416  
; TYPE: PRT  
; ORGANISM: Streptomyces venezuelae  
US-09-836-821-39

Query Match 16.8%; Score 365.5; DB 10; Length 416;  
Best Local Similarity 28.3%; Pred. No. 6.2e-28;  
Matches 97; Conservative 62; Mismatches 163; Indels 21; Gaps 7;

QY 76 EDYRHSFSECPPIPREAGEAYDFTPTSDPPEQRFALANQVGMVVDKLENRIQELA 135  
DB 72 KQWR--NSTTPTLEAEALNNHML--DPPHTRKRLVAREFTMRVVELLRPRVQEV 127

QY 136 CSLIESL--RPQGCNFTEDYAEPPPIRIFMLLAGLPEEDIPHLKYLTDQMTDPGSMTF 193  
DB 128 DGLVDMLAAPDGRADLMESLAWPLITVISSELLGVPEPDRAAFRVWIDAFVFPDPAQA 187

QY 194 ABAEALYDYLIPIEQRQKPGTDAIS--IVANGVNGRPITSDKAKMCGLLVGGT 252  
DB 188 QTAMAEMSGYLSRLIDSKRGQDGEDLLSALVTSDEDSRLTSEELLGMAHILLVAGHET 247

QY 253 VVNFLSFMEFLAKSPEHRQELIERPERIPACEELLRRFSLVADGRILTSDFE----- 307  
DB 248 TVNLITANGYALLSHPDQALRADMTLLDGAVEMLR-----YEGPVESATYRFPV 302

QY 308 --HGVLKKGQDQILLPQMLSGLDERKNACPMHVDFSRQKVSHTTTFHGSHLCLGQHLARR 365

Sequence 13, Application US/10201365  
Publication No. US20030148469A1  
GENERAL INFORMATION:  
APPLICANT: ASHLEY, Gary  
APPLICANT: BETLACH, Melanie C.  
APPLICANT: BETLACH, Mary  
APPLICANT: MCDANIEL, Robert  
APPLICANT: TANG, Li  
TITLE OF INVENTION: COMBINATORIAL POLYKETIDE LIBRARIES PRODUCED USING A MODULAR  
TITLE OF INVENTION: PKS GENE CLUSTER AS SCAFFOLD  
FILE REFERENCE: 300622002103  
CURRENT APPLICATION NUMBER: US/10/201,365  
CURRENT FILING DATE: 2002-07-22  
PRIOR APPLICATION NUMBER: US 09/142,908  
PRIOR FILING DATE: 1998-03-28  
PRIOR APPLICATION NUMBER: US 09/073,538  
PRIOR FILING DATE: 1998-05-06  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: Patent in Ver. 2.0  
SEQ ID NO 13  
LENGTH: 416  
TYPE: PRT  
ORGANISM: Streptomyces venezuelae  
US-10-201-365-13

Query Match 16.8%; Score 365.5; DB 14; Length 416;  
Best Local Similarity 28.3%; Pred. No. 6.2e-28;  
Matches 97; Conservative 62; Mismatches 163; Indels 21; Gaps 7;

QY 76 EDYRHFSECCFIPREAGEAYDFITSDPPQRCFRALANQVGMVVDKLENRIQELA 135  
Db 72 KQWR--NSTTPTLTAEEAALNNHLES--DPPHTLRLKLVAREFTMRRAVELLRPRVQELV 127  
QY 136 CSLIESL--RPQGCNFTEDYAEPPPIRIFMLLAGLPEDIPHLKYLTDQMTDPGSMTF 193  
Db 128 DGLVDMLAAPDGRADLMESLAWPLTIVISELLGVPEPDRAAFRVMTDAFVFPDDPAQA 187  
QY 194 ABAKEALVDYLPIITEQRQKPGTDAIS-IVANGQVNGRPITSDEAKRMCGLLLVGLDT 252  
Db 188 QTAAEMSGYLSRLIDSKRGQDGEDLLSALVTSDEDSRLTSEELGMAHILLVAGHET 247  
QY 253 VVNFLSFMELAKSPHQRQELIERPERIPAAACEILLRRLSVADGRILTSDYEF----- 307  
Db 248 TVNLIANGMYALLSHFDQLAALRADMTLLDGAVEMLR-----YEGPVESATYRFPVEPV 302  
QY 308 --HGVLKKGQILLPQMLSGLDERKNACPMFVDFSRQKVSHTTTFHSHLCLGQHILARR 365  
Db 303 DLDGTVIPAGDTLVVLADAHRTPERFPDPRHFRDIRDTAGHLAFGHGIFHCIGAPLARI 362  
QY 366 ELIVTLKZWLTIPDFS--IAPGAQIOHKSGLVSGVQALPLVW 406  
Db 363 EAR-AVRALLERCPLDLDVSEGLVWVYVNPMPMRIGKALPIRW 405

Search completed: April 6, 2004, 19:14:30  
JOB time : 35.974 secs

RESULT 15  
US-10-539-18  
Sequence 18, Application US/10160539  
Publication No. US20030162262A1  
GENERAL INFORMATION:  
APPLICANT: ASHLEY, Gary  
APPLICANT: BETLACH, Melanie C.  
APPLICANT: BETLACH, Mary C.  
APPLICANT: MCDANIEL, Robert  
APPLICANT: TANG, Li  
TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE  
FILE REFERENCE: 300622002120  
CURRENT APPLICATION NUMBER: US/10/160,539  
CURRENT FILING DATE: 2002-05-29  
PRIOR APPLICATION NUMBER: US/09/657,440  
PRIOR FILING DATE: 2000-09-07  
PRIOR APPLICATION NUMBER: 09/320,878  
PRIOR FILING DATE: 1999-05-27  
PRIOR APPLICATION NUMBER: CIP OF 09/141,908



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 6, 2004, 18:49:16 ; Search time 14.9588 Seconds  
(without alignments)  
1428.803 Million cell updates/sec

Title: US-09-246-451A-11

Perfect score: 2180

Sequence: 1 TTTETQSNANLAPLPPHPE.....IVSGVQALPLVWDPATTKAV 414

Scoring table: BLOSUM62

Gapop 13.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 339414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/prodata/2/iaa/5A.COMB.pep.\*
- 2: /cgn2\_6/prodata/2/iaa/5B.COMB.pep.\*
- 3: /cgn2\_6/prodata/2/iaa/6A.COMB.pep.\*
- 4: /cgn2\_6/prodata/2/iaa/6B.COMB.pep.\*
- 5: /cgn2\_6/prodata/2/iaa/PCTUS.COMB.pep.\*
- 6: /cgn2\_6/prodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	365.5	16.8	416	3	US-09-320-878-18
2	365.5	16.8	416	3	US-09-105-537-39
3	365.5	16.8	416	4	US-09-141-908-13
4	365.5	16.8	416	4	US-09-557-440-18
5	349	16.0	437	4	US-09-252-991A-17836
6	340	15.6	403	6	5212296-5
7	333	15.3	409	3	US-09-385-028-12
8	333	15.3	409	4	US-09-726-614-12
9	333	15.3	409	4	US-09-385-040-12
10	330	15.1	406	6	5212296-6
11	324.5	14.9	399	4	US-08-765-907A-10
12	316	14.5	412	1	US-08-102-863-11
13	316	14.5	412	5	PCT-US92-10885-11
14	271.5	12.5	419	3	US-09-335-409-8
15	271.5	12.5	419	3	US-09-413-814-71
16	271.5	12.5	419	4	US-09-568-102-8
17	271.5	12.5	419	4	US-09-567-969-8
18	271.5	12.5	419	4	US-09-568-480-8
19	271.5	12.5	419	4	US-09-568-486-8
20	271.5	12.5	419	4	US-09-568-472-8
21	271.5	12.5	419	4	US-09-567-899-8
22	261	12.0	395	4	US-09-262-965-129
23	234	10.7	468	4	US-09-252-991A-32437
24	168.5	7.7	189	4	US-09-679-279-20
25	160	7.3	422	2	US-09-096-982-5
26	160	7.3	422	2	US-08-653-650A-5
27	160	7.3	474	2	US-09-096-982-8

28	160	7.3	474	2	US-08-653-650A-8	Sequence 8, Appli
29	157	7.2	443	2	US-09-096-982-9	Sequence 9, Appli
30	157	7.2	443	2	US-08-653-650A-9	Sequence 9, Appli
31	155	7.1	422	1	US-08-396-218-2	Sequence 2, Appli
32	155	7.1	422	1	US-08-760-116-2	Sequence 2, Appli
33	130	6.0	512	2	US-08-194-981B-5	Sequence 5, Appli
34	127	5.8	516	4	US-09-215-694-16	Sequence 16, Appli
35	123.5	5.7	382	3	US-09-320-878-7	Sequence 7, Appli
36	123.5	5.7	382	4	US-09-141-908-7	Sequence 7, Appli
37	123.5	5.7	382	4	US-09-657-440-7	Sequence 22, Appli
38	123.5	5.7	402	3	US-09-105-537-22	Sequence 2, Appli
39	123.5	5.7	503	4	US-09-583-447A-2	Sequence 4, Appli
40	123.5	5.7	3782	3	US-09-105-537-4	Sequence 4, Appli
41	123	5.6	504	4	US-09-583-447A-4	Sequence 4, Appli
42	122	5.6	503	4	US-09-144-167-2	Sequence 24, Appli
43	121.5	5.6	524	4	US-09-126-420A-24	Sequence 6, Appli
44	118	5.4	513	3	US-08-948-564-6	Sequence 3, Appli
45	113.5	5.2	490	1	US-08-201-118-3	Sequence 3, Appli

#### ALIGNMENTS

RESULT 1  
US-09-320-878-18  
; Sequence 18, Application US/09320878A  
; Patent No. 6117659  
; GENERAL INFORMATION:  
; APPLICANT: ASHLEY, Gary  
; APPLICANT: BETLACH, Melanie C.  
; APPLICANT: BETLACH, Mary C.  
; APPLICANT: MCDANIEL, Robert  
; APPLICANT: TANG, Li  
; TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE  
; FILE REFERENCE: 300622002120  
; CURRENT APPLICATION NUMBER: US/09/320,878A  
; CURRENT FILING DATE: 1999-05-27  
; EARLIER APPLICATION NUMBER: CIP OF 09/141,908  
; EARLIER FILING DATE: 1998-08-28  
; EARLIER APPLICATION NUMBER: CIP OF 09/073,538  
; EARLIER FILING DATE: 1998-05-06  
; EARLIER APPLICATION NUMBER: CIP OF 08/846,247  
; EARLIER FILING DATE: 1997-04-30  
; EARLIER APPLICATION NUMBER: 60/119,139  
; EARLIER FILING DATE: 1999-02-08  
; EARLIER APPLICATION NUMBER: 60/100,880  
; EARLIER FILING DATE: 1998-09-22  
; EARLIER APPLICATION NUMBER: 60/087,080  
; EARLIER FILING DATE: 1998-05-28  
; NUMBER OF SEQ ID NOS: 34  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 18  
; LENGTH: 416  
; TYPE: PRT  
; ORGANISM: Streptomyces venezuelae  
US-09-320-878-18

Query Match 16.8%; Score 365.5; DB 3; Length 416;  
Best Local Similarity 28.3%; Pred. No. 2.1e-29;  
Matches 97; Conservative 62; Mismatches 163; Indels 21; Gaps 7;

Qy	76	EDVPHSQCFFIPREAGEAYDFIPTSDPPEQRQFRALANQVVMVVDKLENRIOELA	135
Db	72	KDWR--NSTPLTEAEALNNHML--DPRHTRLKLVAAREFTMRVELLRPRVQIV	127
Qy	136	CSLIESL--RPQGCNFTEDYAEPPFIRIPLMLAGLPEEDIPHLKYLTDQMTDQSGMTF	193
Db	128	DGLVDMALAAPDGRADLMESLAWPLTIVISELLGVPEPDRAAFRVMTDAFVFPDPAQA	187
Qy	194	AEAKEALYDLIPIEIQROKPGTDAIS--IVANQVNGRPITSDKEMCGELLVGGSDT	252
Db	188	QTAMAENSGYLSK--LDSKRGQDGEDLLSALVRTSDEDSRLTSBELLMALILLVAGHET	247

QY 253 VNFSLFSMEFLAKSPHRLQELIERPERIPAAACEBLLRRFSLVADGRILTSDYEF----- 307  
 Db 248 TVNLIAANGMYALLSHPDQALRADMTLGDGAVEEMLR-----YEGPVESATYRFPPEVP 302  
 QY 308 --HGVCLKKGQDILLPQMLSGDLERKKNACPMHVDPSRQKVSHTTFHGSHLCLGQHLLARR 365  
 Db 303 DLDTGTVIPAGDTVLVLADAHRTPEFDPHREDIRDTAGHLAFGHHGHCIGAPLARL 362  
 QY 366 EIIVTLKEMWTRIPDFS--IAPGAIQHKSGIVSGVQALPLVW 406  
 Db 363 EARIAVALLERCPLDALDVSPGELVWYFNPIMIRGLKALPIRW 405

## RESULT 2

US-09-105-537-39  
 ; Sequence 39, Application US/09105537A  
 ; Patent No. 6265202  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Sherman, D.H.  
 ; APPLICANT: Liu, H.  
 ; APPLICANT: Xue, Y.  
 ; APPLICANT: Zhao, L.  
 ; TITLE OF INVENTION: DNA encoding methymycin and pikromycin  
 ; FILE REFERENCE: 600.438US1  
 ; CURRENT APPLICATION NUMBER: US/09/105.537A  
 ; CURRENT FILING DATE: 1998-06-26  
 ; NUMBER OF SEQ ID NOS: 43  
 ; SOFTWARE: FASTSEQ for Windows Version 3.0  
 ; SEQ ID NO 39  
 ; LENGTH: 416  
 ; TYPE: PRT  
 ; ORGANISM: Streptomyces venezuelae  
 US-09-105-537-39

Query Match 16.8%; Score 365.5; DB 3; Length 416;  
 Best Local Similarity 28.3%; Pred. No. 2.1e-29;  
 Matches 97; Conservative 62; Mismatches 163; Indels 21; Gaps 7;

QY 76 EDYRHSSECFPTPREAGEAYDEFTSMDBPEQRFALANQVVGMPVVDKLENRIQELA 135  
 Db 72 KDWK--NSTTTLTEAEAAALNNMLES--DPPRHLRLKLVAREFTMRVRLPRVQEI 127  
 QY 136 CSLTESL--RPQGCNFTEDYAEFPPIRIFMLLAGLPEEDIPHLKYLTDQMTDPGSMTF 193  
 Db 128 DGLVDMLAAPDGRADLMESLAWPLPITVSELLGVPEPDAFRVMTDAFVDDPAQA 187  
 QY 194 AEAKALYDYLPIIEORRQKPGTDAIS--IVANGQVNGRPITSDEAKMCGLLLVGGDLT 252  
 Db 188 QTAWAEMSGVLSRLIDSKRGQDGEDLLSALVTSDEDSRLTSBELLAGMAHILLVAGHET 247  
 QY 253 VNFSLFSMEFLAKSPHRLQELIERPERIPAAACEBLLRRFSLVADGRILTSDYEF----- 307  
 Db 248 TVNLIAANGMYALLSHPDQALRADMTLGDGAVEEMLR-----YEGPVESATYRFPPEVP 302  
 QY 308 --HGVCLKKGQDILLPQMLSGDLERKKNACPMHVDPSRQKVSHTTFHGSHLCLGQHLLARR 365  
 Db 303 DLDTGTVIPAGDTVLVLADAHRTPEFDPHREDIRDTAGHLAFGHHGHCIGAPLARL 362  
 QY 366 EIIVTLKEMWTRIPDFS--IAPGAIQHKSGIVSGVQALPLVW 406  
 Db 363 EARIAVALLERCPLDALDVSPGELVWYFNPIMIRGLKALPIRW 405

## RESULT 3

US-09-141-908-13  
 ; Sequence 13, Application US/09141908  
 ; Patent No. 6503741  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ashley, Gary  
 ; APPLICANT: Betlach, Melanie C.  
 ; APPLICANT: Mcdaniel, Robert  
 ; APPLICANT: Tang, Li

; TITLE OF INVENTION: Combinatorial Polyketide Libraries Produced Using a  
 ; TITLE OF INVENTION: Modular PKS Gene Cluster as Scaffold  
 ; FILE REFERENCE: 300622002100  
 ; CURRENT APPLICATION NUMBER: US/09/141.908  
 ; CURRENT FILING DATE: 1998-08-28  
 ; EARLIER APPLICATION NUMBER: CIP OF 09/073,538  
 ; EARLIER FILING DATE: 1998-05-06  
 ; EARLIER APPLICATION NUMBER: CIP OF 08/846,247  
 ; EARLIER FILING DATE: 1997-04-30  
 ; EARLIER APPLICATION NUMBER: PROV. 60/076,919  
 ; EARLIER FILING DATE: 1998-03-05  
 ; EARLIER APPLICATION NUMBER: PROV. 60/087,080  
 ; EARLIER FILING DATE: 1998-05-28  
 ; NUMBER OF SEQ ID NOS: 31  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 13  
 ; LENGTH: 416  
 ; TYPE: PRT  
 ; ORGANISM: Streptomyces venezuelae  
 US-09-141-908-13

Query Match 16.8%; Score 365.5; DB 4; Length 416;  
 Best Local Similarity 28.3%; Pred. No. 2.1e-29;  
 Matches 97; Conservative 62; Mismatches 163; Indels 21; Gaps 7;

QY 76 EDYRHSSECFPTPREAGEAYDEFTSMDBPEQRFALANQVVGMPVVDKLENRIQELA 135  
 Db 72 KDWK--NSTTTLTEAEAAALNNMLES--DPPRHLRLKLVAREFTMRVRLPRVQEI 127  
 QY 136 CSLTESL--RPQGCNFTEDYAEFPPIRIFMLLAGLPEEDIPHLKYLTDQMTDPGSMTF 193  
 Db 128 DGLVDMLAAPDGRADLMESLAWPLPITVSELLGVPEPDAFRVMTDAFVDDPAQA 187  
 QY 194 AEAKALYDYLPIIEORRQKPGTDAIS--IVANGQVNGRPITSDEAKMCGLLLVGGDLT 252  
 Db 188 QTAWAEMSGVLSRLIDSKRGQDGEDLLSALVTSDEDSRLTSBELLAGMAHILLVAGHET 247  
 QY 253 VNFSLFSMEFLAKSPHRLQELIERPERIPAAACEBLLRRFSLVADGRILTSDYEF----- 307  
 Db 248 TVNLIAANGMYALLSHPDQALRADMTLGDGAVEEMLR-----YEGPVESATYRFPPEVP 302  
 QY 308 --HGVCLKKGQDILLPQMLSGDLERKKNACPMHVDPSRQKVSHTTFHGSHLCLGQHLLARR 365  
 Db 303 DLDTGTVIPAGDTVLVLADAHRTPEFDPHREDIRDTAGHLAFGHHGHCIGAPLARL 362  
 QY 366 EIIVTLKEMWTRIPDFS--IAPGAIQHKSGIVSGVQALPLVW 406  
 Db 363 EARIAVALLERCPLDALDVSPGELVWYFNPIMIRGLKALPIRW 405

## RESULT 4

US-09-657-440-18  
 ; Sequence 18, Application US/09657440  
 ; Patent No. 6509455  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ashley, Gary  
 ; APPLICANT: Betlach, Melanie C.  
 ; APPLICANT: Betlach, Mary C.  
 ; APPLICANT: Mcdaniel, Robert  
 ; APPLICANT: Tang, Li  
 ; TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE  
 ; FILE REFERENCE: 300622002120  
 ; CURRENT APPLICATION NUMBER: US/09/657,440  
 ; CURRENT FILING DATE: 2000-09-07  
 ; PRIOR APPLICATION NUMBER: 09/320,878  
 ; PRIOR FILING DATE: 1999-05-27  
 ; PRIOR APPLICATION NUMBER: CIP OF 09/141,908  
 ; PRIOR FILING DATE: 1998-08-28  
 ; NUMBER OF SEQ ID NOS: 34  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 18  
 ; LENGTH: 416  
 ; TYPE: PRT

```

Query Match      15.6%; Score 340; DB 6; Length 403;
Best Local Similarity 28.9%; Pred. No. 9.3e-27;
Matches 105; Conservative 63; Mismatches 165; Indels 30; Gaps 13

QY      63 WIATRGQLIREAVEDYRHFSEC-----PFIPREAGEAYDFIPT--SMDPPQROQFPALA 115
DB      52 WIVTRHQDVRAVLGDP- FSDAHTGTGFPF.LTAGGREIIIGTNPTELRMDPPEHARLRML 110

QY      116 NQVGMVPVVKLENRIQELACSLIESLRP-QGQCNFTEDYAEPPFIRIPMLLAGLPSEDI 174
DB      111 TADFIVKVKVAMRPEVQRLADLDVDRMTTGTTSADLVTFAFLPESLVITCLLLGVPEYDH 170

QY      175 PHL----KYLTQVTRPDGSMTPAEAKALDYILPIIEQRKQKPGCTDAIS-IVANGQVN 229
DB      171 AFQERSRVLLTLESTPE-----EVRAAQDELLLEYLARLARTKXRPDDAIIISRVARGELD 227

QY      230 GRPITSDEAKRMCGLLIVGGLDTVVNFISFSGMEF.LAKSPEHROELIERPERIPAAACELL 289
DB      228 DTQIAT-----MGRLLIVAGHETTANMTALSTLVLRNPDQLARUAEFALVKGAVEELL 282

QY      290 RRSLVADG--RIILTSVEFHGVOLKKGDOILLPOMLSGLDERKNACP--MHVDFSRQKV 345
DB      283 RYLTIWNGVPRIATEDVLIGRTIAGEGVLC--YISANDRAEVFPQGDLDLDVARDAR 340

QY      346 SHITTFHGSHLCQGHILARRRIIVTLKEMWLRIPDFSIA-POAQIOHKSGLI-VSGVOALP 423
DB      341 RHVAFGFGVHCQCLQOPLARVELQAIATETLLRRLPDLRLAVPHEEIPFPRGDMIAIVGHSIP 400

QY      404 LVW 406
DB      401 IAW 403

RESULT 7
US-09-385-028-12
; Sequence 12, Application US/09385028
; Segment No. 6232106
; Patent No. 6232106
; GENERAL INFORMATION:

```

APPLICANT: Susan E. Jensen  
APPLICANT: Kwamena A. Aidoo  
APPLICANT: Ashish S. Paradkar  
TITLE OF INVENTION: DNA Sequence Encoding Enzymes of Clavulanic  
Patent No. 6232106  
TITLE OF INVENTION: Acid Biosynthesis  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: JACOBSON, PRICE, HOLMAN & STERN, PLLC  
STREET: The Jenner Building, 400 Seventh Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30 (EPC)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/385,028  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/790,462  
FILING DATE: 29-JAN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: D. Douglas Price  
REGISTRATION NUMBER: 24,514  
REFERENCE/DOCKET NUMBER: 1418/P57452US2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 638-6666  
TELEFAX: (202) 39305350  
TELEX: RCA 248593 IDEA UR  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 409 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-385-028-12

Query Match 15.3%; Score 333; DB 3; Length 409;  
Best Local Similarity 29.0%; Pred. No. 5.1e-26;  
Matches 122; Conservative 60; Mismatches 185; Indels 54; Gaps 18;  
QY 16 PHVPEHLVDFDMNPSNLSAGVQEAVALQESNVPLDVTWCNGGH-WIATRGQLIREA 74  
DB 14 PAYPMHRYCPVD---PPQLAGLSRQKAASRVT-----LW---DGSQVWLVTSHAGARAV 62  
QY 75 YEDYRHFS-SECPRTIP-----REAGEAYDF-PTSMPPPEQRFRA-----LANQV 118  
DB 63 LGDRFRFVAVTAPGFPMLTRTSQLVANPESASF--RMDDPQHSRLSMLTRDFLARRA 120  
QY 119 VGM-PVVDKLENRIQELACSLIESLRPOGQCNFTEDYAEPPPIRIFMLLAGLPEEDIHPL 177  
DB 121 EALRPVAVREL---LDEILGLLVKGERP---VDLVAGLTIPVPSRVITLLFGAGDDRRRFI 174  
QY 178 K-----YLTQMTDTRDGSMTFAEKEALDYLIPIEQRQKPGTDAISVANGVNGRPI 233  
DB 175 EDRSAVILDRGYTPE---QVAKARDELGVYRELVEERIEPNPTDLISRLVIDQVRPGHL 231  
QY 234 TSDEAKRMCGLLLVGGLDTVNVNFSFMEFLAKSPEHEQELIERPERIPAAECLELLRRFS 293  
DB 232 RVEEMVPMCRLLLVAGHGTTTSQASLSLLTDFELAGRTEDFALLPKAVEELLRRHS 291  
QY 294 LVADG---RILTSDEFFHGVQLKGDQILLPQMLSGLDERKACPMHVDPSKQVSHHTFG 351  
DB 292 IVONGLAARAAVEDVQLDDVLIRAGRGVVLISAGNRJSTVFPDPDVRDVRDARRHLAFG 351  
QY 352 HGSLLCLOHAR---RIIIVTLKWLTRIPDFSIA-PCAQIQKSGIVS-GVQMLPLVW 406

DB 352 HGMHQCQLQWLARVLELEILAAVLWRM---PGARLAVPFEELDFRHEVSSYGLGALPVTW 408  
QY 407 D 407  
DB 409 Z 409  
RESULT 8  
US-09-726-614-12  
Sequence 12, Application US/09726614  
Patent No. 6514735  
GENERAL INFORMATION:  
APPLICANT: Susan E. Jensen  
APPLICANT: Kwamena A. Aidoo  
APPLICANT: Ashish S. Paradkar  
TITLE OF INVENTION: DNA Sequence Encoding Enzymes of Clavulanic  
Patent No. 6514735  
TITLE OF INVENTION: Acid Biosynthesis  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: JACOBSON, PRICE, HOLMAN & STERN, PLLC  
STREET: The Jenner Building, 400 Seventh Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/726,614  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/790,462  
FILING DATE: 29-JAN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: D. Douglas Price  
REGISTRATION NUMBER: 24,514  
REFERENCE/DOCKET NUMBER: 1418/P57452US2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 638-6666  
TELEFAX: (202) 39305350  
TELEX: RCA 248593 IDEA UR  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 409 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-726-614-12

Query Match 15.3%; Score 333; DB 4; Length 409;  
Best Local Similarity 29.0%; Pred. No. 5.1e-26;  
Matches 122; Conservative 60; Mismatches 185; Indels 54; Gaps 18;  
QY 16 PHVPEHLVDFDMNPSNLSAGVQEAVALQESNVPLDVTWCNGGH-WIATRGQLIREA 74  
DB 14 PAYPMHRYCPVD---PPQLAGLSRQKAASRVT-----LW---DGSQVWLVTSHAGARAV 62  
QY 75 YEDYRHFS-SECPRTIP-----REAGEAYDF-PTSMPPPEQRFRA-----LANQV 118  
DB 63 LGDRFRFVAVTAPGFPMLTRTSQLVANPESASF--RMDDPQHSRLSMLTRDFLARRA 120  
QY 119 VGM-PVVDKLENRIQELACSLIESLRPOGQCNFTEDYAEPPPIRIFMLLAGLPEEDIHPL 177  
DB 121 EALRPVAVREL---LDEILGLLVKGERP---VDLVAGLTIPVPSRVITLLFGAGDDRRRFI 174  
QY 178 K-----YLTQMTDTRDGSMTFAEKEALDYLIPIEQRQKPGTDAISVANGVNGRPI 233

Db 175 EDRSAVLIDRGTYPE---QVAKARDELVDGYLRELVEREINENPGTDLISRLVIDQVRPHL 231  
 QY 234 TSDPAKMGCLLVGLDVTWNFLSFMFELAKSPEHQELIERPERIPAAACEILLRFS 293  
 Db 232 RVEEMVENCRLLVAGHGTTSQASLSLLITDPELAGRLTEDPALLPKAVEILLRFS 291  
 QY 294 LVADG--RLTSDYEFHGVOLKGDQIL--PQMLSGDLERKNACPMHVDPSQKVSHTTEG 351  
 Db 292 IVQNGLARAAVEDVOLDLVIRAGEGVVLSLSAGNRDETFPDPDRVDVDRARHLAFG 351  
 QY 352 HGSCLCGLQHLAR---REIIVTLKEMLTRIPDFSIA-PGAQIOHKSGIVS-GVQALPLVW 406  
 Db 352 HGMHCCGQLARVLELILAAVLRW---PGARLAVPFEELDFRHEVSSYGLGALPVTW 408  
 QY 407 D 407  
 Db 409 Z 409

RESULT 9  
 US-09-385-040-12  
 ; Sequence 12, Application US/09385040  
 ; Patent No. 6589775  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Jensen, Susan E  
 ; APPLICANT: Aideo, Kwamena A  
 ; APPLICANT: Paraskar, Ashish S  
 ; TITLE OF INVENTION: DNA SEQUENCE ENCODING ENZYMES OF CLAVULANIC ACID  
 ; FILE REFERENCE: 09/385,040  
 ; CURRENT APPLICATION NUMBER: US/09/385,040  
 ; PRIOR FILING DATE: 1999-08-30  
 ; PRIOR APPLICATION NUMBER: US 08/790,462  
 ; PRIOR FILING DATE: 1997-01-29  
 ; NUMBER OF SEQ ID NOS: 25  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 12  
 ; LENGTH: 409  
 ; TYPE: PRT  
 ; ORGANISM: Streptomyces clavuligerus  
 US-09-385-040-12

Query Match 15.3%; Score 333; DB 4; Length 409;  
 Best Local Similarity 29.0%; Pred. No. 5.1e-26;  
 Matches 122; Conservative 60; Mismatches 185; Indels 54; Gaps 18;  
 QY 16 PHVPEHLVDFEDMNPMSLQAGVQAWVLQSNVPDLVWTRCNGH-WIATRGQLIREA 74  
 Db 14 PAYPMHRVCPVD---PPQLAGLSQKAASRYT---LW---DGSQVWLVTSHAGARAV 62  
 QY 75 YEDYRHFS-SECPFIP-----REAGEAYDFIPTSMDPPEQCFRA-----LANQV 118  
 Db 63 LGDRFTAVTSAPGPFMLTRTSQLVANPESASF1--RMDDFQHSRLRSLMTRDFIARRA 120  
 QY 119 VGM-PWVDKLENRIQELACSTESLRPQCQNFTEIYAEPPFIRIFMLLAGPEEDIPHL 177  
 Db 121 EALRPVREL---LDELGLLVKGERP---VDLVAGLTIPVPSRVITLFGAGDORREFI 174  
 QY 178 K----YLTQWTRPGSMTFAKALYDYLIPILQEQKPGTDAISIVANGQVNGRPI 233  
 Db 175 EDRSAVLIDRGTYPE---QVAKARDELVDGYLRELVEERINENPGTDLISRLVIDQVRPHL 231  
 QY 234 TSDPAKMGCLLVGLDVTWNFLSFMFELAKSPEHQELIERPERIPAAACEILLRFS 293  
 Db 232 RVEEMVENCRLLVAGHGTTSQASLSLLITDPELAGRLTEDPALLPKAVEILLRFS 291  
 QY 294 LVADG--RLTSDYEFHGVOLKGDQIL--PQMLSGDLERKNACPMHVDPSQKVSHTTEG 351  
 Db 292 IVQNGLARAAVEDVOLDLVIRAGEGVVLSLSAGNRDETFPDPDRVDVDRARHLAFG 351  
 QY 352 HGSCLCGLQHLAR---REIIVTLKEMLTRIPDFSIA-PGAQIOHKSGIVS-GVQALPLVW 406  
 Db 352 HGMHCCGQLARVLELILAAVLRW---PGARLAVPFEELDFRHEVSSYGLGALPVTW 408

QY 407 D 407  
 Db 409 Z 409  
 RESULT 10  
 5212296-6  
 ; Patent No. 5212296  
 ; APPLICANT: DEAN, CAROLINE; HARDER, PATRICIA A.; LETO, KENNETH  
 ; J.; O'KEEFE, DANIEL P.; OMER, CHARLES A.; ROMESSER, JAMES A.  
 ; TEPPERMAN, JAMES M.  
 ; TITLE OF INVENTION: EXPRESSION OF HERBICIDE METABOLIZING  
 ; CYTOCHROMES  
 ; NUMBER OF SEQUENCES: 19  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/07/569,781  
 ; FILING DATE: 23-AUG-1990  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 464,499  
 ; FILING DATE: 12-JAN-1990  
 ; APPLICATION NUMBER: 405,605  
 ; FILING DATE: 11-SEP-1989  
 ; SEQ ID NO: 6  
 ; LENGTH: 406  
 5212296-6  
 Query Match 15.1%; Score 330; DB 6; Length 406;  
 Best Local Similarity 26.3%; Pred. No. 1e-25;  
 Matches 104; Conservative 74; Mismatches 193; Indels 24; Gaps 12;  
 QY 31 PSNLSAGVQ--EAWAVLQESNVPLVWTRCNGH-WIATRGQLIREAYEDYR----- 79  
 Db 17 PSNRSCPVLDPGYAQLRDTPCGLHRVTLYDGRQAWVTKEAARKLLGDPRLSSNRTDD 76  
 QY 80 HFSSECPFIP--REAGEAYDFIPTSMDPPEQCFRANALANQVGMPPVDKLENRIQELACS 137  
 Db 77 NFPATSPRFAVRSPQAF---IGLDPEHGTRRRMTISEFTVKIKGMRPEVEEVVHG 132  
 QY 138 LIESLRPQ--QCNTFEDVAEPEFIRIFMLLAGPEEDIPHLKYLTDQMTPEQSSMTFAEA 196  
 Db 133 FIDEMLAGFTADLVSQFALPVPSMVICRLGLVPYADHFFQDASKRLVQSTDRQASALTA 192  
 QY 197 KEALYDYLIPILTEQRQKPGTDAI--SIVANGQVNGKPIITSDEAKRMCGLLVGLDVTVN 255  
 Db 193 ENDLAGILDCLITQCTEFGAGLVGALVADQLANGE-IDREELISTAMLLIAGHETTAS 251  
 QY 256 FLSFSEMF-AKSPHQRQELIERPERIPAAACEILLRFSL--VADGRILTSDYEFHGVOLK 313  
 Db 252 MTSLSVITLDDHPEQVAAALRADRSIVPGAVEILLRYLAIDAGGRVATADIEVEGELIR 311  
 QY 314 KEDQILLPQMLSGDLERKNACPMHVDPSQKVSHTTEFGHSHLCLGQHLARREIIVTKE 373  
 Db 312 AGEGVIVVNSIARNGTGVYEDPDALDIHRSARHLLAFGFGVHOCGLQNLARULEVLNA 371  
 QY 374 WLTRIPDFSIA-PGAQIOHKSG-IVSGVQALPLVW 406  
 Db 372 IMDRPTLELVPEQLVLRPGTTIQGVNELPVTW 406

RESULT 11  
 US-08-765-907A-10  
 ; Sequence 10, Application JS/08765907A  
 ; Patent No. 6352839  
 ; GENERAL INFORMATION:  
 ; APPLICANT: BLANC, Veronique  
 ; APPLICANT: THIBAUT, Denis  
 ; APPLICANT: BAWAS-JACQUES, Nathalie  
 ; APPLICANT: BLANCHE, Francis  
 ; APPLICANT: COUZET, Joel  
 ; APPLICANT: BARRIERE, Jean-Claude  
 ; APPLICANT: DEBUSSCHE, Laurent  
 ; APPLICANT: FAMECHON, Alain

APPLICANT: PARIS, Jean-Marc  
APPLICANT: DUTRUC-ROSSET, Gilles  
TITLE OF INVENTION: Streptogramins And Method For Preparing Same By  
TITLE OF INVENTION: Mutasynthesis  
FILE REFERENCE: Streptogramin genes  
CURRENT APPLICATION NUMBER: US/08/765,907A  
CURRENT FILING DATE: 1997-03-20  
NUMBER OF SEQ ID NOS: 17  
SOFTWARE: Patent In Ver. 2.0  
SEQ ID NO 12  
LENGTH: 399  
TYPE: PRT  
ORGANISM: Streptomyces pristinaespiralis  
US-08-765-907A-10

Query Match 14.9%; Score 324.5; DB 4; Length 399;  
Best Local Similarity 29.0%; Pred. No. 3.8e-25;  
Matches 106; Conservative 59; Mismatches 159; Indels 41; Gaps 12;

QY 74 AYEDYRH-----FSSECPFIPREAGEAYDPTSMDDPPEQKQFRALANQVGM 121  
DB 36 AHWFRHADVLTVASDFGVSSQLRLRPGSQALSEQLLSVIDPPMHTLRLVSAFT 95  
QY 122 PVVDKLENRQELACSLIESLRPGQC-NFTEDVAEPFIRIFMLLAGLPEEDIPHLKYL 180  
DB 96 RTVADLEPRVTELAGQLDAV--DGDYFDLWADPAYPLVIVIAELLGVPPADRTLSFW 153  
QY 181 TDQWTR-----PDGSMYFAEAKALYDYLPIIEQRQKPGTDAISIVA 224  
DB 154 SDRMLQWQVADDMQGDGDAEDYQRLVKEPMAHAYLHDHVDTRARPANDLISLV 213  
QY 225 NGQVNGRPITSDAKRMCGLLVGLSDTVNFLSFMFLAKSPHRLQELIERPER--IP 282  
DB 214 AARVEGRLEDEQIVERGALLMAGHVSSTMLGNVTVLCKDHP--RAEAAARADRSIP 271  
QY 283 AACBELLR-RFSLWADGRITSDYEFHGVQLKGGQILLQMLSL-GLDERKXNACPMHVD 340  
DB 272 ALIEEVLELRPPITVMARVTTKDTVLAGTTIPAG-RMVVPSLLSANDEQVETDPDHL 330  
QY 341 SRQKVSHTTGHSHLCLGQHLARREIIVTLKWLTRIPDPSIAPGACIO-HKSGIVSGV 399  
DB 331 AREG-RQIAFGHGIHYCLGAPLARLEGRINALEALDFRDPFSDTGAKIRYHRDGLF-GV 388  
QY 400 QALPL 404  
DB 389 KNLPL 393

## RESULT 12

US-08-102-863-11  
Sequence 11, Application US/08102863  
Patent No. 5466590  
GENERAL INFORMATION:  
APPLICANT: SARIASLANI, SIMA  
TITLE OF INVENTION: CONSTITUTIVE  
TITLE OF INVENTION: EXPRESSION OF P450SOY  
TITLE OF INVENTION: AND FERREDOXIN-SOY IN  
TITLE OF INVENTION: STREPTOMYCES  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: E. I. DU PONT DE NEMOURS  
ADDRESSEE: AND COMPANY  
STREET: 1007 MARKET STREET  
CITY: WILMINGTON  
STATE: DELAWARE  
COUNTRY: USA  
ZIP: 19898  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0,  
SOFTWARE: Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/102,863  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/807,001  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: GALLEGOS, R. THOMAS  
REGISTRATION NUMBER: 32,692  
REFERENCE/DOCKET NUMBER: CR-9000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 302-892-7342  
TELEFAX: 302-892-7949  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 412 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-102-863-11

Query Match 14.5%; Score 316; DB 1; Length 412;

Best Local Similarity 27.2%; Pred. No. 3.1e-24;

Matches 84; Conservative 56; Mismatches 163; Indels 6; Gaps 5;  
QY 103 MDPEQKQFRALANQVGMVVDKLENRQELACSLIESLRPGQ-QCNFTEDVAEPFIR 161  
DB 105 VDDPEHNTQRMILPTFSVKRIGALRPRIQTVDRLLDMERQGPAPBLVSAFALPVFSM 164  
QY 162 IFMLLAGLPEEDIPHLKYLTDQMTTRPDGSMYFAEAKALYDYLPIIEQRQKPGTDAIS 221  
DB 165 VICALLGVVADHAPFEERSQRLLRGPGADVDNRARQDELYLGALIDRKAEPDGLD 224  
QY 222 IVANGVNGRPITSDAKRMCGLLVGLSDTVNFLSFMFLAKSPHRLQELIERPERI 281  
DB 225 ELTHRDHPDGPVDRQELVAFVITLLIAGHETTANMISLGTFTLLSHPEQLAALRAGTST 284  
QY 282 PAACEELLRRFSLWADG--RLTSDYEFHGVQLKGGQILLQMLSGLDERKXNACPMHVD 339  
DB 285 AVVVEELL-RFLSIAEGQLRATEDMEVDGATIRKGGVWFSTSLINKDADVFPRAETLD 343  
QY 340 FSROKVSHTTGHSHLCLGQHLARREIIVTLKWLTRIPDPSIA-PGAIQHKSG-IVS 397  
DB 344 WDRPARHHLAFGFGVHQCGLQNLARAEALDIAMRTLFERLPGRLAVPAHEIRHKPGDTIQ 403  
QY 398 GVQALPLVW 406  
DB 404 GLLDLPVAV 412

## RESULT 13

PCT-US92-10885-11  
Sequence 11, Application PC/TUS9210885  
GENERAL INFORMATION:  
APPLICANT: SARIASLANI, SIMA  
TITLE OF INVENTION: CONSTITUTIVE  
TITLE OF INVENTION: EXPRESSION OF P450SOY  
TITLE OF INVENTION: AND FERREDOXIN-SOY IN  
TITLE OF INVENTION: STREPTOMYCES  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: E. I. DU PONT DE NEMOURS  
ADDRESSEE: AND COMPANY  
STREET: 1007 MARKET STREET  
CITY: WILMINGTON  
STATE: DELAWARE  
COUNTRY: USA  
ZIP: 19898  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch,  
MEDIUM TYPE: 1.0 MB

```

;
; COMPUTER: Macintosh
; OPERATING SYSTEM: Macintosh System, 6.0
; SOFTWARE: Microsoft Word, 4.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/10885
; FILING DATE: 19921216
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: GALLEGO, R. THOMAS
; REGISTRATION NUMBER: 32,692
; REFERENCE/DOCKET NUMBER: CR-9000-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-892-7342
; TELEFAX: 302-892-7949
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 412 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; PCT-US92-10885-11

```

```

Query Match      14.5%; Score 316; DB 5; Length 412;
Best Local Similarity 27.2%; Pred. No. 3.1e-24;
Matches 84; Conservative 56; Mismatches 163; Indels 6; Gaps 5;

```

```

QY 103 MDPBQRQFALANQVGMVVDKLENIQELACSLIESLRPQG-QCNFTEDYAEPPPIR 161
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ 105 VDEPHNTQRMILPTFSVRKIGALRPRIQETVDRLIDAMERQPPRAELVSFAFALPVPSM 164
QY 162 IFVLLAGLPEEDIPHLKYLTDQTRPDGSMTEAEAKALDYLIPIIQEQRQKPGTDAIS 221
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 165 VICALLGVFADHAFPERQQLRGPGADQVNPARDLEEYLGLIDRRAREPGGGLD 224
QY 222 VANGQVNGRPITSDAKRKGILLVGLDITVNVFLSFEELAKSPHEHQELIERPERI 281
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ 225 ELIHRDHPDQVDEQVAFVILLIAGHTTANWISLGTFTLLSHPEQLAALRAGGTST 284
QY 282 PAACEELLRFSLVADG--RILTSDFEFGVQLKKGQIQLPQMLGLDERKACPMHVD 339
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ 285 AVVVEELL-RFLSIAEGLQRLATEDMEVDGATIRKGGVVFSTSLINRDADVFPPAETLD 343
QY 340 FSRQKVSHTTFHGSHLCLGQHLARREIIVTLKELTRIPDFSTA-PGAQIQHKSQ-IVS 397
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 344 WDPARHHLAFGVGHQCLQNLARAEIDIAMRTFLERPLGLRVAVPAHELRKPGDTIQ 403
QY 398 GVOQLPLVW 406
Db      : : : : :
DQ 404 GLLDLPVAV 412

```

```

RESULT 14
US-09-335-409-8
; Sequence 8, Application: US/09335409
; Patent No. 6121329
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/335,409
; CURRENT FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Sorangium cellulosum

```

```

US-09-335-409-8
Query Match      12.5%; Score 271.5; DB 3; Length 419;
Best Local Similarity 23.8%; Pred. No. 1.4e-19;
Matches 100; Conservative 67; Mismatches 179; Indels 75; Gaps 14;

QY 6 QSNANLAPLPHVPEHLVCEDMYNPSNLSAGVQE---AWAVLQESNVVDLVWTRCNGS 61
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 3 QEQANQSETKP-----AFDPPFAP-----GYAEDPPAERLUREA-TPIFYWD--EGR 48
QY 62 HWIATRGQ-----LIREAYETVRFHSECEPPIPREAGEAYDFITSDPPQQR 109
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 49 SWLTRYHDYSAVFRDERFVSRSESSAIP-----ELSDMKKYGLGFLPPEDHA 104
QY 110 QFRALANQVGMVVDKLENIQELACSLIESLRPQGQCNFTEDYAEPPPIRIFMLLAGL 169
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 105 FVRKLVNPSFTSRAIDLLRAEIQRTVDQLDARSQGEFFVDVRYAEGIPRAISALLKV 164
QY 170 PEEDIPHLKYLTDQTRPDGSMTEAEAKALDYLIPI-----I 207
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 165 PAE-----CDEKFRFGSAT-----ARALGVGLVPOVDEETKTLVASVTEGLALLHDV 212
QY 208 IEORRQKP-GTDAISIVANGQVNGRPITSDAKRKGILLVGLDITVNVFLSFEELAK 266
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 213 LDERRRNPLENDVLTMLLQAEADGSRSLTKELVALVGAIIAAGTTTIVLIAFAVLNLR 272
QY 267 SPEHQELIERPERIPACBELLRFSLVADG--RILTSDFEFGVQLKKGQI--LLPQ 322
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 273 SPEALELVKAEPPGLMRNALDEVLRDNILRIGTVRFARQDLEYCGASIKKGMVELLPS 332
QY 323 MLSGLDERKACPMHVDVFSRQKVSHTTFHGSHLCLGQHLARREIIVTLKELTRIPDFS 382
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 333 ALR--DGVTSRPRVDFVRDRTGSLAYGRPHVCPGVSLARLAREIAVGIFRFPENK 390
QY 383 I 383
Db      :
QY 391 I 391

```

```

RESULT 15
US-09-413-814-71
; Sequence 71, Application US/09413814
; Patent No. 6225064
; GENERAL INFORMATION:
; APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH
; APPLICANT: Bristol-Myers Squibb, Co.
; APPLICANT: Beyer, Stefan
; APPLICANT: Bloeker, Helmut
; APPLICANT: Brandt, Petra
; APPLICANT: Cino, Paul M
; APPLICANT: Dougherty, Brian A
; APPLICANT: Goldberg, Steven L
; APPLICANT: Hoeffle, Gerhard
; APPLICANT: Mueller, Joachim
; APPLICANT: Reichenbach, Hans
; TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polyketide or
; TITLE OF INVENTION: heteropolyketide compounds
; FILE REFERENCE: PCT/US 99/23535
; CURRENT APPLICATION NUMBER: US/09/413,814
; CURRENT FILING DATE: 1999-10-07
; EARLIER APPLICATION NUMBER: DE 198 46 493.2
; EARLIER FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 71
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-09-413-814-71

```

```

Query Match      12.5%; Score 271.5; DB 3; Length 419;
Best Local Similarity 23.8%; Pred. No. 1.4e-19;
Matches 100; Conservative 67; Mismatches 179; Indels 75; Gaps 14;

```





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 5, 2004, 18:53:57 ; Search time 34.974 Seconds  
(without alignments)  
3108.883 Million cell updates/sec

Title: US-09-246-451a-12

Perfect score: 2179

Sequence: 1 TTTTQSNANLAPLPVHVE.....IVSGVQALPLVWDPTTKAV 414

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1071772 seqs, 262633353 residues

Total number of hits satisfying chosen parameters: 1071772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 5%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2179	100.0	414	15	US-10-453-104-12
2	2173	99.7	414	15	US-10-453-104-11
3	2169	99.5	414	15	US-10-453-104-2
4	2162	99.2	414	15	US-10-453-104-13
5	383.5	17.6	404	14	US-10-214-446-50
6	382.5	17.6	404	14	US-10-214-446-40
7	378	17.3	399	14	US-10-314-657-37
8	373	17.1	416	14	US-10-156-761-14997
9	371.5	17.0	416	9	US-09-861-289-39
10	371.5	17.0	416	9	US-09-860-846-39
11	371.5	17.0	416	10	US-09-988-384B-39
12	371.5	17.0	416	10	US-09-836-821-39
13	371.5	17.0	416	10	US-09-793-708-18
14	371.5	17.0	416	14	US-10-201-365-13
15	371.5	17.0	416	14	US-10-160-539-18

Sequence 39, Appl  
Sequence 8376, Ap  
Sequence 9525, Ap  
Sequence 38, Appl  
Sequence 20, Appl  
Sequence 9914, Ap  
Sequence 11073, A  
Sequence 559, App  
Sequence 4, Appli  
Sequence 12, Appl  
Sequence 32, Appl  
Sequence 13776, A  
Sequence 2, Appli  
Sequence 16, Appl  
Sequence 14, Appl  
Sequence 14659, A  
Sequence 8710, Ap  
Sequence 36, Appl  
Sequence 9703, Ap  
Sequence 22, Appl  
Sequence 8, Appli  
Sequence 10431, A  
Sequence 22, Appl  
Sequence 6, Appli  
Sequence 62, Appl  
Sequence 4117, Ap  
Sequence 7959, Ap  
Sequence 14, Appl  
Sequence 46, Appl

#### ALIGNMENTS

RESULT 1  
US-10-453-104-12  
; Sequence 12, Application US/10453104  
; Publication No. US20030207345A1  
; GENERAL INFORMATION:  
; APPLICANT: California Institute of Technology;  
; APPLICANT: Frances H. Arnold  
; APPLICANT: Hyun Joo  
; TITLE OF INVENTION: Oxygenase Enzymes and Screening Method  
; FILE REFERENCE: 4058/1E827-US3  
; CURRENT APPLICATION NUMBER: US/10/453.104  
; CURRENT FILING DATE: 2003-06-02  
; PRIOR APPLICATION NUMBER: US 09/661,093  
; PRIOR FILING DATE: 2000-09-13  
; PRIOR APPLICATION NUMBER: US 09/246,451  
; PRIOR FILING DATE: 1999-02-09  
; PRIOR APPLICATION NUMBER: US 60/094,403  
; PRIOR FILING DATE: 1998-07-28  
; PRIOR APPLICATION NUMBER: US 60/106,840  
; PRIOR FILING DATE: 1998-11-03  
; PRIOR APPLICATION NUMBER: US 60/086,206  
; PRIOR FILING DATE: 1998-05-21  
; PRIOR APPLICATION NUMBER: US 60/106,834  
; PRIOR FILING DATE: 1998-11-03  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 12  
; LENGTH: 414  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; OTHER INFORMATION: Mutant M7-6H  
US-10-453-104-12

Query Match 100.0%; Score 2179; DB 15; Length 414;  
Best Local Similarity 100.0%; Pred. No. 6.5e-212;  
Matches 414; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 TTETIQSNANLAPLPHVPEHLVDFDMYNPSNLISAGVQEAVALQESNVPDLVWTRCNG 60
Db 1 TTETIQSNANLAPLPHVPEHLVDFDMYNPSNLISAGVQEAVALQESNVPDLVWTRCNG 60
Qy 61 GHWIATRGQILREAYEDYRHFSSCPPIPREAGEAYDFIPTSMPPQORFRALANQVWG 120
Db 61 GHWIATRGQILREAYEDYRHFSSCPPIPREAGEAYDFIPTSMPPQORFRALANQVWG 120
Qy 121 MPVVDKLENRIQELACSLIESLRPQGCNFTEDYAEFPPIRIFMLLAGLPEDIPHLKYL 180
Db 121 MPVVDKLENRIQELACSLIESLRPQGCNFTEDYAEFPPIRIFMLLAGLPEDIPHLKYL 180
Qy 181 TDQWTRPDGSMTPAEAKEALYDYLPIIEORRQKPGTDAISIVANGVNGRPITSDAKR 240
Db 181 TDQWTRPDGSMTPAEAKEALYDYLPIIEORRQKPGTDAISIVANGVNGRPITSDAKR 240
Qy 241 MCGLLVGGGLDVTWNFLSFMELAKSPERQELIERPELIPAAACELLRRFSIVADGRI 300
Db 241 MCGLLVGGGLDVTWNFLSFMELAKSPERQELIERPELIPAAACELLRRFSIVADGRI 300
Qy 301 LTSDFEFGVOLKKGQOILLPQMLSGLDERKNACPMHVDFSRQKVSHTTFEGSHLCLGQ 360
Db 301 LTSDFEFGVOLKKGQOILLPQMLSGLDERKNACPMHVDFSRQKVSHTTFEGSHLCLGQ 360
Qy 361 HLAARREIIVTLKEWLTRIPDFSAPGAQIQHKSIVSGVQALPLVMDPATTKAV 414
Db 361 HLAARREIIVTLKEWLTRIPDFSAPGAQIQHKSIVSGVQALPLVMDPATTKAV 414

```

## RESULT 2

```

US-10-453-104-11
; Sequence 11, Application US/10453104
; Publication No. US20030207345A1
; GENERAL INFORMATION:
; APPLICANT: California Institute of Technology;
; APPLICANT: Frances H. Arnold
; APPLICANT: Hyun Joo
; TITLE OF INVENTION: Oxygenase Enzymes and Screening Method
; FILE REFERENCE: 4058/1E827-US3
; CURRENT APPLICATION NUMBER: US/10/453,104
; CURRENT FILING DATE: 2003-06-02
; PRIOR APPLICATION NUMBER: US 09/661,093
; PRIOR FILING DATE: 2000-09-13
; PRIOR APPLICATION NUMBER: US 09/246,451
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: US 60/094,403
; PRIOR FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: US 60/106,840
; PRIOR FILING DATE: 1998-11-03
; PRIOR APPLICATION NUMBER: US 60/086,206
; PRIOR FILING DATE: 1998-05-21
; PRIOR APPLICATION NUMBER: US 60/106,834
; PRIOR FILING DATE: 1998-11-03
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Fast-Seq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mutant M7-4H
US-10-453-104-11

```

```

Query Match 99.7%; Score 2173; DB 15; Length 414;
Best Local Similarity 99.8%; Pred. No. 2.6e-211;
Matches 413; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy 1 TTETIQSNANLAPLPHVPEHLVDFDMYNPSNLISAGVQEAVALQESNVPDLVWTRCNG 60
Db 1 TTETIQSNANLAPLPHVPEHLVDFDMYNPSNLISAGVQEAVALQESNVPDLVWTRCNG 60
Qy 61 GHWIATRGQILREAYEDYRHFSSCPPIPREAGEAYDFIPTSMPPQORFRALANQVWG 120

```

```

Db 61 GHWIATRGQILREAYEDYRHFSSCPPIPREAGEAYDFIPTSMPPQORFRALANQVWG 120
Qy 121 MPVVDKLENRIQELACSLIESLRPQGCNFTEDYAEFPPIRIFMLLAGLPEDIPHLKYL 180
Db 121 MPVVDKLENRIQELACSLIESLRPQGCNFTEDYAEFPPIRIFMLLAGLPEDIPHLKYL 180
Qy 181 TDQWTRPDGSMTPAEAKEALYDYLPIIEORRQKPGTDAISIVANGVNGRPITSDAKR 240
Db 181 TDQWTRPDGSMTPAEAKEALYDYLPIIEORRQKPGTDAISIVANGVNGRPITSDAKR 240
Qy 241 MCGLLVGGGLDVTWNFLSFMELAKSPERQELIERPELIPAAACELLRRFSIVADGRI 300
Db 241 MCGLLVGGGLDVTWNFLSFMELAKSPERQELIERPELIPAAACELLRRFSIVADGRI 300
Qy 301 LTSDFEFGVOLKKGQOILLPQMLSGLDERKNACPMHVDFSRQKVSHTTFEGSHLCLGQ 360
Db 301 LTSDFEFGVOLKKGQOILLPQMLSGLDERKNACPMHVDFSRQKVSHTTFEGSHLCLGQ 360
Qy 361 HLAARREIIVTLKEWLTRIPDFSAPGAQIQHKSIVSGVQALPLVMDPATTKAV 414
Db 361 HLAARREIIVTLKEWLTRIPDFSAPGAQIQHKSIVSGVQALPLVMDPATTKAV 414

```

## RESULT 3

```

US-10-453-104-2
; Sequence 2, Application US/10453104
; Publication No. US20030207345A1
; GENERAL INFORMATION:
; APPLICANT: California Institute of Technology;
; APPLICANT: Frances H. Arnold
; APPLICANT: Hyun Joo
; TITLE OF INVENTION: Oxygenase Enzymes and Screening Method
; FILE REFERENCE: 4058/1E827-US3
; CURRENT APPLICATION NUMBER: US/10/453,104
; CURRENT FILING DATE: 2003-06-02
; PRIOR APPLICATION NUMBER: US 09/661,093
; PRIOR FILING DATE: 2000-09-13
; PRIOR APPLICATION NUMBER: US 09/246,451
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: US 60/094,403
; PRIOR FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: US 60/106,840
; PRIOR FILING DATE: 1998-11-03
; PRIOR APPLICATION NUMBER: US 60/086,206
; PRIOR FILING DATE: 1998-05-21
; PRIOR APPLICATION NUMBER: US 60/106,834
; PRIOR FILING DATE: 1998-11-03
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Fast-Seq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 414
; TYPE: PRT
; ORGANISM: P. Putida
US-10-453-104-2

```

```

Query Match 99.5%; Score 2169; DB 15; Length 414;
Best Local Similarity 99.5%; Pred. No. 6.7e-211;
Matches 412; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy 1 TTETIQSNANLAPLPHVPEHLVDFDMYNPSNLISAGVQEAVALQESNVPDLVWTRCNG 60
Db 1 TTETIQSNANLAPLPHVPEHLVDFDMYNPSNLISAGVQEAVALQESNVPDLVWTRCNG 60
Qy 61 GHWIATRGQILREAYEDYRHFSSCPPIPREAGEAYDFIPTSMPPQORFRALANQVWG 120
Db 61 GHWIATRGQILREAYEDYRHFSSCPPIPREAGEAYDFIPTSMPPQORFRALANQVWG 120
Qy 121 MPVVDKLENRIQELACSLIESLRPQGCNFTEDYAEFPPIRIFMLLAGLPEDIPHLKYL 180
Db 121 MPVVDKLENRIQELACSLIESLRPQGCNFTEDYAEFPPIRIFMLLAGLPEDIPHLKYL 180
Qy 181 TDQWTRPDGSMTPAEAKEALYDYLPIIEORRQKPGTDAISIVANGVNGRPITSDAKR 240

```

```

Db 181 TDQMPDGSMTFAEAKAALDYLLPIIEQRKQKPGCTDAISIVANGQVNGRPITSDEAKR 240
QY 241 MCGLLLVGGGLDVTNNFLSFMFLAKSPHRRQELIERPELIPACBELLRRFSLVADGRI 300
Db 241 MCGLLLVGGGLDVTNNFLSFMFLAKSPHRRQELIERPELIPACBELLRRFSLVADGRI 300
QY 301 LTSDYEFHGVQVKKGQDQILLPQMLSGLDREKRNACPMHVDPSRKQVSHHTTFGHSHLCIGQ 360
Db 301 LTSDYEFHGVQVKKGQDQILLPQMLSGLDREKRNACPMHVDPSRKQVSHHTTFGHSHLCIGQ 360
QY 361 HLAARREIIVTLKEWLTRIPDFSIAPGAQIQHKSIGVSGVQALPLVMDPATTKAV 414
Db 361 HLAARREIIVTLKEWLTRIPDFSIAPGAQIQHKSIGVSGVQALPLVMDPATTKAV 414

RESULT 4
US-10-453-104-13
; Sequence 13, Application US/10453104
; Publication No. US20030207345A1
; GENERAL INFORMATION:
; APPLICANT: California Institute of Technology;
; APPLICANT: Frances H. Arnold
; APPLICANT: Hyun Joo
; TITLE OF INVENTION: Oxygenase Enzymes and Screening Method
; FILE REFERENCE: 4C58/1E827-US3
; CURRENT APPLICATION NUMBER: US/10/453,104
; CURRENT FILING DATE: 2003-06-02
; PRIOR APPLICATION NUMBER: US 09/661,093
; PRIOR FILING DATE: 2000-09-23
; PRIOR APPLICATION NUMBER: US 09/246,451
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: US 60/094,403
; PRIOR FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: US 60/106,840
; PRIOR FILING DATE: 1998-11-03
; PRIOR APPLICATION NUMBER: US 60/086,206
; PRIOR FILING DATE: 1998-05-21
; PRIOR APPLICATION NUMBER: US 60/106,834
; PRIOR FILING DATE: 1998-11-03
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mutant M7-8H
US-10-453-104-13

```

```

Query Match 99.2%; Score 2162; DB 15; Length 414;
Best Local Similarity 99.5%; Pred. No. 3.5e-210;
Matches 412; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTETIQSNANLAPLPHVPEHLVDFEDMYPNSLSAGVQAWAVLQESNVPDLVWTRCNG 60
Db 1 TTETIQSNANLAPLPHVPEHLVDFEDMYPNSLSAGVQAWAVLQESNVPDLVWTRCNG 60
QY 61 GHWIATRGQILREAYEDYRHFSSECFPIPEAGEAYDPIPTSMCPQORFALANQVVG 120
Db 61 GHWIATRGQILREAYEDYRHFSSECFPIPEAGEAYDPIPTSMCPQORFALANQVVG 120
QY 121 MPVVDKLENRIQELACSLIESLRPQOCNTEYARPPRIRIFMLLAGLPEEDIPLHKYL 180
Db 121 MPVVDKLENRIQELACSLIESLRPQOCNTEYARPPRIRIFMLLAGLPEEDIPLHKYL 180
QY 181 TDQMPDGSMTFAEAKAALDYLLPIIEQRKQKPGCTDAISIVANGQVNGRPITSDEAKR 240
Db 181 TDQMPDGSMTFAEAKAALDYLLPIIEQRKQKPGCTDAISIVANGQVNGRPITSDEAKR 240
QY 241 MCGLLLVGGGLDVTNNFLSFMFLAKSPHRRQELIERPELIPACBELLRRFSLVADGRI 300
Db 241 MCGLLLVGGGLDVTNNFLSFMFLAKSPHRRQELIERPELIPACBELLRRFSLVADGRI 300

```

```

QY 301 LTSDYEFHGVQVKKGQDQILLPQMLSGLDREKRNACPMHVDPSRKQVSHHTTFGHSHLCIGQ 360
Db 301 LTSDYEFHGVQVKKGQDQILLPQMLSGLDREKRNACPMHVDPSRKQVSHHTTFGHSHLCIGQ 360
QY 361 HLAARREIIVTLKEWLTRIPDFSIAPGAQIQHKSIGVSGVQALPLVMDPATTKAV 414
Db 361 HLAARREIIVTLKEWLTRIPDFSIAPGAQIQHKSIGVSGVQALPLVMDPATTKAV 414

RESULT 5
US-10-214-446-50
; Sequence 50, Application US/10214446
; Publication No. US20030180742A1
; GENERAL INFORMATION:
; APPLICANT: Weiner, David
; APPLICANT: Burk, Mark J.
; APPLICANT: Hitchman, Tim
; APPLICANT: Pujol, Catherine
; APPLICANT: Richardson, Toby
; APPLICANT: Short, Jay M.
; TITLE OF INVENTION: F450 ENZYMES, NUCLEIC ACIDS ENCODING
; TITLE OF INVENTION: THEM AND METHODS OF MAKING AND USING THEM
; FILE REFERENCE: 09010-500001
; CURRENT APPLICATION NUMBER: US/10/214,446
; CURRENT FILING DATE: 2002-08-05
; PRIOR APPLICATION NUMBER: US 60/309,497
; PRIOR FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 50
; LENGTH: 404
; TYPE: PRT
; ORGANISM: Bacterial
US-10-214-446-50

```

```

Query Match 17.6%; Score 383.5; DB 14; Length 404;
Best Local Similarity 27.9%; Pred. No. 8.4e-30;
Matches 98; Conservative 64; Mismatches 156; Indels 33; Gaps 5;

QY 60 GHWIATRGQILREAYEDYRHFSSECFPIPEAGEAYDPIPTSM----- 103
Db 32 GRTWFLPHHADIRTLALRDERFSAS-----RKAGGFVNQPPAEVRPEFAFNEAISRWIV 85
QY 104 --DPBQORFALANOVGVVVKLENRIQELACSLIESLRPQOCNTEYARPPRIR 161
Db 86 LHDQPEHRQRLQMQQGFTRRLITTEPKIORVCDLIDAFVKRGSSTEFMTETIAHPFAK 145
QY 162 IFMLLAGLPEEDIPLHKYLTDQMPDGSMTFAEAKAALDYLLPIIEQRKQ 213
Db 146 VIAEMLVNPEDYPAFVWSEDLINPAGSLRPTLEMFRAAQDGLAMMDYFARLLPERRE 205
QY 214 KPQTDALISIVANGQVNGRPITSDEAKMCGLLLVGGGLDVTNNFLSFMFLAKSPHROE 273
Db 206 NPGDGLVSLLSAESGEWMTABQVLANCQTIIVAGHETTRNLVANGVELLRYPBQCAL 265
QY 274 LIERPELIPACBELLRRFSLVADGRILLSDYEFHGVQVKKGQDQILLPQMLSGLDREK 332
Db 266 LERPELIPACBELLRRFSLVADGRILLSDYEFHGVQVKKGQDQILLPQMLSGLDREK 325
QY 333 ACMPHVDPSRKQVSHHTTFGHSHLCIGQHLARREIIVTLKEWLTRIPDFS 383
Db 326 DDPDTDLTENTPGHLAGFGWPEVCVGAALAELEGQVSFRTLLDRLPGLEL 376

```

```

RESULT 6
US-10-214-446-43
; Sequence 40, Application US/10214446
; Publication No. US20030180742A1
; GENERAL INFORMATION:
; APPLICANT: Weiner, David
; APPLICANT: Burk, Mark J.
; APPLICANT: Hitchman, Tim
; APPLICANT: Pujol, Catherine

```

```

; APPLICANT: Richardson, Toby
; APPLICANT: Short, Jay M.
; TITLE OF INVENTION: P450 ENZYMES, NUCLEIC ACIDS ENCODING
; FILE REFERENCE: 09010-500001
; CURRENT APPLICATION NUMBER: US/10/214.446
; PRIOR FILING DATE: 2002-08-05
; PRIOR FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 40
; LENGTH: 404
; TYPE: PRT
; ORGANISM: Bacterial
; ORGANISM: Streptomyces atroolivaceus
US-10-214-446-40

Query Match      17.6%; Score 382.5; DB 14; Length 404;
Best Local Similarity 30.3%; Pred. No. 1.1e-29;
Matches 120; Conservative 58; Mismatches 187; Indels 31; Gaps 15;

QY 31 PSNLSAGVQAMAVLQESNVDPDLVWTRCNGGH--WIATRGQILREAYEDYRHFSSECPFI 88
Db 20 PFNEAGISLADAYEAREQFGLRVEMAYGEPALATR-----YADARLVLGDRRFS 72
QY 89 -----PREAGEAYDFTPTSDPPEQRFALANQVGVVVDKLENRIQELACSLI 139
Db 73 RAEGARDEPRQSGRDSGLSDPDPHTRLTLVAKAFTHQVKEKLRPAVRLADELI 132
QY 140 ESLRPGQ-QCNFTEDYAEPPPIRIFMLLAGLPEEDI PHLKYLTDQMTDPGSMT---FAE 195
Db 133 DKVATGAPVDLVEEPALPVGVVICQLGVVEEVRFRFAWS-D-AALSTSLTAEEFDA 191
QY 196 AKCALVDYLPILIQRQKPGTDAISIVANGQVNGRPTTDEAKRMCGLLVGGLOTVNV 255
Db 192 NQSELRAYMGLIEDHRAHPREDLITGLIEARDDDRLTQELVDLCVGLVAGHETAT 251
QY 256 FLFSFSEFLAKSPHQRQELTERPELIPAAACBELLRSELSVADG---RLTSDYEPHGVQ 311
Db 252 QIFNVVTLDRPEQWNLREDPELVPTAVEELM-RFVPLGSGASFRYATEDVEVGGTL 310
QY 312 LKXGQDILLPQMLSG-LDERKNACPMHVDFFGRQKVSHTTFGGHSHLCGLQHLARREIIVTL 371
Db 311 VRAGEPVLVAVGAANRDPARFAPQELDLAREGNQHLGFGHGVHHCILGAPLARLEQAL 370
QY 372 KEMTRIPDSIAPGAQIQHKS-GIVSGVQALPLVW 406
Db 371 GALLRLPGURIA--GDIEWKQMLVGRPTLPVGVW 404

RESULT 7
US-10-314-657-37
; Sequence 37, Application US/10314657
; Publication No. US20030175888A1
; GENERAL INFORMATION:
; APPLICANT: SHEN, Ben
; APPLICANT: CHENG, Yi-Qiang
; APPLICANT: TANG, Gong-Li
; TITLE OF INVENTION: Discrete Acyltransferases Associated with Type I Polyketide
; FILE REFERENCE: 054030-3021
; CURRENT APPLICATION NUMBER: US/10/314.657
; CURRENT FILING DATE: 2002-12-09
; PRIOR FILING DATE: 2002-03-22
; PRIOR FILING DATE: 2002-03-22
; PRIOR FILING DATE: 2001-03-26
; NUMBER OF SEQ ID NOS: 214
; SOFTWARE: PatentIt version 3.2
; SEQ ID NO 37
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Streptomyces atroolivaceus

```

US-10-314-657-37

```

Query Match      17.3%; Score 378; DB 14; Length 399;
Best Local Similarity 28.3%; Pred. No. 3e-29;
Matches 111; Conservative 70; Mismatches 175; Indels 36; Gaps 9;

QY 38 VQEAVALQESNVDPDLVWTRCNGGH--ATRGQILREAYEDYRHFSSECPFI PREAGE--- 94
Db 21 IHKPAELRETDPLARVLPYGGEGKWYTR-----YDDVRAANSDFRFSQAQIGETPT 73
QY 95 -----AYDFTPTSDPPEQRFALANQVGVVVDKLENRIQELACSLIESLRPGQ 147
Db 74 RITPLARRSDTI-LSLDPPEHTRLRLSKAFTARRMGAMQSWLEELFAGLLDGVERTGH 132
QY 148 -CNFTEDYAEPPPIRIFMLLAGLPEEDI PHLKYLTDQMTDPGSMTFAAKE-----A 199
Db 133 PADIVRLAQFPTTAVICRLGVVEYEDRGRFQHWSEVI-----MSTTAYSKEEAVSADAS 187
QY 200 LYDLILPIIEQRQKPGTDAISIVANGQVNGRPTTDEAKRMCGLLVGGLOTVNVFLSF 259
Db 188 IRVLADIVSARRAAPHDDLGLVLSARDDDRLTEJELITFGVTLVAGHETSAHQNGN 247
QY 260 SMEFLAKSPHQRQELTERPELIPAAACBELLRFSL---VADGRILTSYRPHGVOLKGD 316
Db 248 MVYALLTHEDQLSLRQPELLPRAVELLAFVPLGVNGVGNARIALEDELVELSGGTVRAGE 307
QY 317 QILLPQMLSG-LDERKNACPMHVDFFSRQKVSHTTFGGHSHLCGLQHLARREIIVTLKEWLT 376
Db 308 GVVAANVNAANDRPAFDDPDRLEITREKNPHLAFGHGAHYCLGAGLAQMLARVLAIGGLLE 367
QY 377 RIPDPSIA-PGAQIQHKS-GIVSGVQALPLVW 406
Db 368 RFPGLRLAVPADQVEMTKTGLFRGPQRLPIAW 399

RESULT 8
US-10-156-761-14997
; Sequence 14997, Application US/10156761
; Publication No. US20030119318A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156.761
; CURRENT FILING DATE: 2002-05-29
; PRIOR FILING DATE: 2001-05-30
; PRIOR FILING DATE: 2001-05-30
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15129
; SEQ ID NO 14997
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-14997

Query Match      17.1%; Score 373; DB 14; Length 416;
Best Local Similarity 29.4%; Pred. No. 1e-28;
Matches 105; Conservative 56; Mismatches 180; Indels 16; Gaps 8;

QY 63 WIATRGQILREAYEDYRHFSSECPFI PR--EAGEAYDFTPT-----SMDPPEQRFAL 114
Db 63 WYVTGHAARALLSDQLSSDRT--LRFPPATTFERFAVTRRVALLGVDDDPHEHTQRM 120
QY 115 ANQVGVVVDKLENRIQELACSLIESLRPGQ-QCNFTEDYAEPPPIRIFMLLAGLPEED 173
Db 121 LVFSFTLKRAAALRPRIQETVVDGLDAMEAQGPPELVSAFALPSPMVICALLGVFPYAD 180

```

174 IPHLYLTQMTDPGSMFAEAKALYDLPIIIBOROKPGTDAISIVANGVNGRPI 233  
181 HFFESQSRRLRGPGLAEVQDRAQDDLYALIDRKREPGDGLDDLIQQLARGTV 240  
234 TDEAKRMCGLLVAGSLCTVNFSLFSMEFLAKSPHQBELIRPELIPAAECLELRRFS 293  
241 DRAELVSLATLLLIAGHETANNISLGTFTLLRHPEQLAELRAEPGLMPAAVEELL-REL 299  
294 LVADG--RLTSDYEHGVLQKEDQILLPQMLSGLDKRNACPMHVDPSRQKSHHTTF 351  
300 STADGHLRVATEDIEVAGTTIRADEGVFATSVINRDAAGFAEPDADLHRSARHHVAFG 359  
352 HGHSLCLOGLARREIIVTKELATRIPOFSI-APGAQIOHKSG-IVSGVQALPLVW 406  
360 FGIHQCLQNLARAEMIALGTLEFLPGLRLAAPADEIPFKGDTIQGMLBPLVTW 416

## RESULT 9

US-09-861-289-39  
Sequence 39, Application US/09861289  
Patent No. US20020112897A1  
GENERAL INFORMATION:  
APPLICANT: Sherman, D.H.  
APPLICANT: Liu, H.  
APPLICANT: Xue, Y.  
APPLICANT: Zhao, L.  
TITLE OF INVENTION: DNA encoding methymycin and pikromycin  
FILE REFERENCE: 600.438US1  
CURRENT APPLICATION NUMBER: US/09/861,289  
PRIOR FILING DATE: 2001-05-18  
PRIOR APPLICATION NUMBER: 09/105,537  
PRIOR FILING DATE: 1998-06-26  
NUMBER OF SEQ ID NOS: 43  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 39  
LENGTH: 416  
TYPE: PRT  
ORGANISM: Streptomyces venezuelae  
US-09-861-289-39

Query Match 17.0%; Score 371.5; DB 9; Length 416;  
Best Local Similarity 28.6%; Pred. No. 1.4e-28;  
Matches 98; Conservative 62; Mismatches 162; Indels 21; Gaps 7;  
QY 76 EDYRHFSSCPFTPREAGAYDFTPSMDPPEQORQFRALANQVVGMPVDKLENRIOELA 135  
DB 72 KDWK--NSTTPTTEAEALNHNMLES--DPRHTRLRKLVARFTMRVVELLRPRVQEI 127  
QY 136 CSLIESL--RPOGOCNFTDYABFPPIRIMLAGLPEEDIPHLKYLTDOMTRPDGSMTF 193  
DB 128 DGLVDAMLAAPDGRADIMESLAWPLPITVISELLGVPEPDRAAFRVWTDAFVFPDPAQA 187  
QY 194 ABAKEALYDLPIIBOROKPGTDAIS-IVANGVNGRPIITSDAKRMCGLLVAGSLGT 252  
DB 188 QTMAEMSGYLSRLIDSKRGQDGEDLLSA-LVTSDEGSLTSEELGMAHILLVAGHET 247  
QY 253 VVNFSLFSMEFLAKSPHQBELIRPELIPAAECLELRRSLVADGRILTSDFEYF----- 307  
DB 248 TVNLIANGMYALLSHPDQALRADMTLLDGAVEEMLR-----YEGVESATYFFPVEPV 302  
QY 308 --HGVOLKKGDDQILLPQMLSGLDKRNACPMHVDPSRQKSHHTTFHGHSLCLGHLARR 365  
DB 303 DLDGTVIPAGTVLVVLADAHRTPERPDPHRRDIRDTAGHLAAGHGHIFCIGAPLARL 362  
QY 366 EIIVTLKEMLTRIPDS--IAPGAQIOHKSGIVSGVQALPLVW 406  
DB 363 EARIAVRALLERCPDLALDVSPGELVWYPNPMIRGLKALPIRW 405

## RESULT 10

US-09-860-846-39  
Sequence 39, Application US/09860846

Patent No. US20020164742A1  
GENERAL INFORMATION:  
APPLICANT: Sherman, D.H.  
APPLICANT: Liu, H.  
APPLICANT: Xue, Y.  
APPLICANT: Zhao, L.  
TITLE OF INVENTION: DNA encoding methymycin and pikromycin  
FILE REFERENCE: 600.438US1  
CURRENT APPLICATION NUMBER: US/09/860,846  
PRIOR FILING DATE: 2001-05-18  
PRIOR APPLICATION NUMBER: 09/105,537  
PRIOR FILING DATE: 1998-06-26  
NUMBER OF SEQ ID NOS: 43  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 39  
LENGTH: 416  
TYPE: PRT  
ORGANISM: Streptomyces venezuelae  
US-09-860-846-39

Query Match 17.0%; Score 371.5; DB 9; Length 416;  
Best Local Similarity 28.6%; Pred. No. 1.4e-28;  
Matches 98; Conservative 62; Mismatches 162; Indels 21; Gaps 7;  
QY 76 EDYRHFSSCPFTPREAGAYDFTPSMDPPEQORQFRALANQVVGMPVDKLENRIOELA 135  
DB 72 KDWK--NSTTPTTEAEALNHNMLES--DPRHTRLRKLVARFTMRVVELLRPRVQEI 127  
QY 136 CSLIESL--RPOGOCNFTDYABFPPIRIMLAGLPEEDIPHLKYLTDOMTRPDGSMTF 193  
DB 128 DGLVDAMLAAPDGRADIMESLAWPLPITVISELLGVPEPDRAAFRVWTDAFVFPDPAQA 187  
QY 194 ABAKEALYDLPIIBOROKPGTDAIS-IVANGVNGRPIITSDAKRMCGLLVAGSLGT 252  
DB 188 QTMAEMSGYLSRLIDSKRGQDGEDLLSA-LVTSDEGSLTSEELGMAHILLVAGHET 247  
QY 253 VVNFSLFSMEFLAKSPHQBELIRPELIPAAECLELRRSLVADGRILTSDFEYF----- 307  
DB 248 TVNLIANGMYALLSHPDQALRADMTLLDGAVEEMLR-----YEGVESATYFFPVEPV 302  
QY 308 --HGVOLKKGDDQILLPQMLSGLDKRNACPMHVDPSRQKSHHTTFHGHSLCLGHLARR 365  
DB 303 DLDGTVIPAGTVLVVLADAHRTPERPDPHRRDIRDTAGHLAAGHGHIFCIGAPLARL 362  
QY 366 EIIVTLKEMLTRIPDS--IAPGAQIOHKSGIVSGVQALPLVW 406  
DB 363 EARIAVRALLERCPDLALDVSPGELVWYPNPMIRGLKALPIRW 405

## RESULT 11

US-09-988-384B-39  
Sequence 39, Application US/0988384B  
Publication No. US20030073824A1  
GENERAL INFORMATION:  
APPLICANT: Sherman, D.H.  
APPLICANT: Liu, H.  
APPLICANT: Xue, Y.  
APPLICANT: Zhao, L.  
TITLE OF INVENTION: DNA encoding methymycin and pikromycin  
FILE REFERENCE: 600.536US1  
CURRENT APPLICATION NUMBER: US/09/988,384B  
CURRENT FILING DATE: 2001-11-19  
PRIOR APPLICATION NUMBER: PCT/US99/14398  
PRIOR FILING DATE: 1999-06-25  
PRIOR APPLICATION NUMBER: US 09/105,537  
PRIOR FILING DATE: 1998-06-26  
NUMBER OF SEQ ID NOS: 53  
SEQ ID NO 39  
LENGTH: 416  
TYPE: PRT  
ORGANISM: Streptomyces venezuelae  
US-09-988-384B-39





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 6, 2004, 13:49:16 ; Search time 14.9588 Seconds  
(without alignments)  
1428.803 Million cell updates/sec

Title: US-09-346-451A-12

Perfect score: 2179

Sequence: 1 TTTCTCSNANLAPLPHVPE.....IVSGVQALPLYWDPNTTKAV 414

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51E25971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

1: /cgn2\_6/prodata/2/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/prodata/2/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/prodata/2/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/prodata/2/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/prodata/2/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/prodata/2/iaa/backfiles1.pep.\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	371.5	17.0	416	3	US-09-320-878-18
2	371.5	17.0	416	3	US-09-105-537-39
3	371.5	17.0	416	4	US-09-141-908-13
4	371.5	17.0	416	4	US-09-657-440-18
5	355	16.3	437	4	US-09-252-991A-17836
6	346	15.9	403	6	5212296-9
7	339	15.6	409	3	US-09-385-028-12
8	339	15.6	409	4	US-09-726-614-12
9	339	15.6	439	4	US-09-385-040-12
10	336	15.4	406	6	5212296-6
11	323.5	14.8	398	4	US-08-765-907A-10
12	315	14.5	412	1	US-08-102-863-11
13	315	14.5	412	5	PCT-US92-10885-11
14	277.5	12.7	419	3	US-09-335-409-8
15	277.5	12.7	419	3	US-09-413-814-71
16	277.5	12.7	419	4	US-09-568-102-8
17	277.5	12.7	419	4	US-09-567-969-8
18	277.5	12.7	419	4	US-09-568-480-8
19	277.5	12.7	419	4	US-09-568-486-8
20	277.5	12.7	419	4	US-09-568-472-8
21	277.5	12.7	419	4	US-09-567-899-8
22	265	12.2	395	4	US-09-266-965-129
23	240	11.0	468	4	US-09-252-991A-32437
24	174.5	7.3	189	4	US-09-679-279-20
25	160	7.3	422	2	US-09-096-982-5
26	160	7.3	422	2	US-08-653-650A-5
27	160	7.3	474	2	US-09-096-982-8

28	160	7.3	474	2	US-08-653-650A-8	Sequence 8, Appli
29	157	7.2	443	2	US-09-096-982-9	Sequence 9, Appli
30	157	7.2	443	2	US-08-653-650A-9	Sequence 9, Appli
31	155	7.1	422	1	US-08-396-218-2	Sequence 2, Appli
32	155	7.1	422	1	US-08-760-116-2	Sequence 2, Appli
33	135	6.2	512	2	US-08-194-981E-5	Sequence 5, Appli
34	129.5	5.9	382	3	US-09-320-878-7	Sequence 7, Appli
35	129.5	5.9	382	4	US-09-141-908-7	Sequence 7, Appli
36	129.5	5.9	382	4	US-09-657-440-7	Sequence 7, Appli
37	129.5	5.9	402	3	US-09-105-537-22	Sequence 22, Appli
38	129.5	5.9	3782	3	US-09-105-537-4	Sequence 4, Appli
39	127.5	5.9	524	4	US-09-126-420A-24	Sequence 24, Appli
40	127	5.8	516	4	US-09-215-694-16	Sequence 16, Appli
41	120.5	5.5	503	4	US-09-583-447A-2	Sequence 2, Appli
42	120	5.5	503	4	US-09-144-367-2	Sequence 2, Appli
43	120	5.5	504	4	US-09-583-447A-4	Sequence 4, Appli
44	119.5	5.5	504	4	US-09-976-594-642	Sequence 642, App
45	119	5.5	501	3	US-08-906-791-2	Sequence 2, Appli

#### ALIGNMENTS

##### RESULT 1

US-09-320-878-18  
; Sequence 18, Application US/09320878A  
; Patent No. 6117659  
; GENERAL INFORMATION:  
; APPLICANT: ASHLEY, Gary  
; APPLICANT: BETLACH, Melanie C.  
; APPLICANT: BETLACH, Mary C.  
; APPLICANT: MC DANIEL, Robert  
; APPLICANT: TANG, Li  
; TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE  
; FILE REFERENCE: 300622002120  
; CURRENT FILING DATE: 1999-05-27  
; EARLIER APPLICATION NUMBER: US/09/320, 878A  
; EARLIER FILING DATE: 1999-05-27  
; EARLIER APPLICATION NUMBER: CIP OF 09/141, 908  
; EARLIER FILING DATE: 1998-08-28  
; EARLIER APPLICATION NUMBER: CIP OF 09/073, 536  
; EARLIER FILING DATE: 1998-05-06  
; EARLIER APPLICATION NUMBER: CIP OF 08/846, 247  
; EARLIER FILING DATE: 1997-04-30  
; EARLIER APPLICATION NUMBER: 60/119, 139  
; EARLIER FILING DATE: 1999-02-08  
; EARLIER APPLICATION NUMBER: 60/100, 880  
; EARLIER FILING DATE: 1998-09-22  
; EARLIER APPLICATION NUMBER: 60/087, 080  
; NUMBER OF SEQ ID NOS: 34  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 18  
; LENGTH: 416  
; TYPE: PRT  
; ORGANISM: Streptomyces venezuelae  
US-09-320-878-18

Qy	76	EDVRFHSSECPPEAGRAYDFIPTSMDPEQOFALANOVVMPVVDKLENRIQELA	135
Db	72	KDWR--NSTPTLTEAALNNHML--DPRTRLRKLVAREFTMRVRLPRVQEV	127
Qy	136	CSLIESL--RPQCCNFTEDYAPFPPIRIFMLAGLPEEDIPHLKYLTDQWTRPDGSMTF	193
Db	128	DGLVDAMLAAPDGRADLMESLAMPITVISELLGVPEPDRAAFRVWTDVFPDDPAQA	187
Qy	194	AEAKELLYLLEIFIROROKPCTDAIS--IVANGQVNGRPITSDAKRMCGLLLVGLDT	252
Db	188	QTAWAEMSGYLSRLIDSKRGQDGLLSALVRTSDEGSLTSELGMHILLVAGHET	247

Query Match 17.0%; Score 371.5; DB 3; Length 416;  
Best Local Similarity 28.6%; Pred. No. 7.7e-31;  
Matches 98; Conservative 62; Mismatches 162; Indels 21; Gaps 7;



```

QY 253 VNFSLFSMEFLAKSPHEHQEIERPELIPAAACEILLRRLSVADGRILTSDYEF----- 307
Db 248 TVNLIANGMYALLSHPDQALRADMTLLDGAVEMLR-----YEGPVESATYRFPVPV 302
QY 308 --HGVLKKGDOILLPQMLSGLDERKNACPMHVDPSRQKVSHTTFGHGSHLCIGQHARR 365
Db 303 DLGGTVIPAGDVTLVVLADAHRTPEFPDHPFRDRTAGHLAFGHGHCIGAPLARL 362
QY 366 EIIVTLEKWLTRIPDFS--IAPGAQIOHKGSGIVSGVQALPLVW 406
Db 363 EARIARALLERCEDLALDVSFGLVWPNMIRGLKALPIRW 405

RESULT 2
US-09-105-537-39
; Sequence 39, Application US/09105537A
; Patent No. 6265202
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600,438USI
; CURRENT APPLICATION NUMBER: US/09/105,537A
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 39
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-105-537-39

```

```

Query Match 17.0%; Score 371.5; DB 3; Length 416;
Best Local Similarity 28.6%; Pred. No. 7.7e-31;
Matches 98; Conservative 62; Mismatches 162; Indels 21; Gaps 7;

QY 76 EDYRHSFSECPFIPREAGEAYDFPTSMPPQROFRALANQVGMVVDKLENRIQELA 135
Db 72 KDMR--NSTTPELTAEEAALNNHLES--DPPHTRLRKLVAAREFTMRVE--LRPRVEIV 127
QY 136 CSLIESL--RPOGQCNFTDYAEPPPIRIFMLAGLPEDIPHLKYITDQWTRPDGSMTF 193
Db 128 DGLVDMLAAPDGRADLMESLAWLPITVISELLGVPPDRAAFVWTDADFVPDDPAQA 187
QY 194 AEAKALYDYLPIIEORROKPGTDAIS--IVANGQVNGRPITTSDEAKRMCGLLVGGLD 252
Db 188 QTAMAEMSGYLSRLIDSKRGQDGEDLLSALVRTSDEGSLTSELLGMHILLVAGHET 247
QY 253 VNFSLFSMEFLAKSPHEHQEIERPELIPAAACEILLRRLSVADGRILTSDYEF----- 307
Db 248 TVNLIANGMYALLSHPDQALRADMTLLDGAVEMLR-----YEGPVESATYRFPVPV 302
QY 308 --HGVLKKGDOILLPQMLSGLDERKNACPMHVDPSRQKVSHTTFGHGSHLCIGQHARR 365
Db 303 DLGGTVIPAGDVTLVVLADAHRTPEFPDHPFRDRTAGHLAFGHGHCIGAPLARL 362
QY 366 EIIVTLEKWLTRIPDFS--IAPGAQIOHKGSGIVSGVQALPLVW 406
Db 363 EARIARALLERCEDLALDVSFGLVWPNMIRGLKALPIRW 405

```

```

RESULT 3
US-09-141-908-13
; Sequence 13, Application US/09141908
; Patent No. 6503741
; GENERAL INFORMATION:
; APPLICANT: ASHLEY, Gary
; APPLICANT: BETLACH, Melanie C.
; APPLICANT: BETLACH, Mary
; APPLICANT: MCDANIEL, Robert
; APPLICANT: TANG, Li

```

```

; TITLE OF INVENTION: Combinatorial Polyketide Libraries Produced Using a
; FILE REFERENCE: Modular PKS Gene Cluster as Scaffold
; CURRENT APPLICATION NUMBER: US/09/141,908
; EARLIER FILING DATE: 1998-08-28
; EARLIER APPLICATION NUMBER: CIP OF 09/073,538
; EARLIER FILING DATE: 1998-05-06
; EARLIER APPLICATION NUMBER: CIP OF 08/846,247
; EARLIER FILING DATE: 1997-04-30
; EARLIER APPLICATION NUMBER: PROV. 60/076,919
; EARLIER FILING DATE: 1998-03-05
; EARLIER APPLICATION NUMBER: PROV. 60/087,080
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 13
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-141-908-13

```

```

Query Match 17.0%; Score 371.5; DB 4; Length 416;
Best Local Similarity 28.6%; Pred. No. 7.7e-31;
Matches 98; Conservative 62; Mismatches 162; Indels 21; Gaps 7;

QY 76 EDYRHSFSECPFIPREAGEAYDFPTSMPPQROFRALANQVGMVVDKLENRIQELA 135
Db 72 KDMR--NSTTPELTAEEAALNNHLES--DPPHTRLRKLVAAREFTMRVE--LRPRVEIV 127
QY 136 CSLIESL--RPOGQCNFTDYAEPPPIRIFMLAGLPEDIPHLKYITDQWTRPDGSMTF 193
Db 128 DGLVDMLAAPDGRADLMESLAWLPITVISELLGVPPDRAAFVWTDADFVPDDPAQA 187
QY 194 AEAKALYDYLPIIEORROKPGTDAIS--IVANGQVNGRPITTSDEAKRMCGLLVGGLD 252
Db 188 QTAMAEMSGYLSRLIDSKRGQDGEDLLSALVRTSDEGSLTSELLGMHILLVAGHET 247
QY 253 VNFSLFSMEFLAKSPHEHQEIERPELIPAAACEILLRRLSVADGRILTSDYEF----- 307
Db 248 TVNLIANGMYALLSHPDQALRADMTLLDGAVEMLR-----YEGPVESATYRFPVPV 302
QY 308 --HGVLKKGDOILLPQMLSGLDERKNACPMHVDPSRQKVSHTTFGHGSHLCIGQHARR 365
Db 303 DLGGTVIPAGDVTLVVLADAHRTPEFPDHPFRDRTAGHLAFGHGHCIGAPLARL 362
QY 366 EIIVTLEKWLTRIPDFS--IAPGAQIOHKGSGIVSGVQALPLVW 406
Db 363 EARIARALLERCEDLALDVSFGLVWPNMIRGLKALPIRW 405

```

```

RESULT 4
US-09-657-440-18
; Sequence 18, Application US/09657440
; Patent No. 6509455
; GENERAL INFORMATION:
; APPLICANT: ASHLEY, Gary
; APPLICANT: BETLACH, Melanie C.
; APPLICANT: BETLACH, Mary C.
; APPLICANT: MCDANIEL, Robert
; APPLICANT: TANG, Li
; TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 300622002120
; CURRENT APPLICATION NUMBER: US/09/657,440
; CURRENT FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: 09/320,878
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: CIP OF 09/141,908
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 18
; LENGTH: 416
; TYPE: PRT

```

ORGANISM: Streptomyces venezuelae

US-09-637-440-18

Query Match 17.0%; Score 37.5; DB 4; Length 416;

Best Local Similarity 28.6%; Pred. No. 7.7e-31;

Matches 98; Conservative 62; Mismatches 162; Indels 21; Gaps 7;

QY 76 EDVRFSSCPFIPRAGEAYDPIPTSMDFPEQORQFRALANOVGMVVDKLENRQELA 135

DB 72 KWR--NSTPTTEAALHNWLES--CPPHTRLRKLVAREFTNRVVELLAPRQEI 127

QY 136 CSLIESL--RPGQCNTEDYAEPPPIRIFMLLAGPEEDIPHLKYLTDQMTFGDSMTF 193

DB 128 DGLVADLAAPDGRADLMESLAWPLITVSELLGVPEPDRAAFRVWTDFAVFPDPAQA 187

QY 194 ABAKALYLYLPIIQRQKPTDAIS--TVANGVNGRPITTSDEAKRMCGLLVGLDT 252

DB 188 QTAMAMSGYLSRLIDSKRGQGEDLLSALVTSDEGSRLTSEELGMHAILLVAGHET 247

QY 253 VNFLSFSEFLAKSPHROELIERPELIPAAACEELLRRFSLVADGRILTSDFEYF--- 307

DB 248 TYNLIANGYALLSHPDQAAIRADMTLLDGAVEMLR-----YEGPVESATYRFPV 302

QY 308 --HGVQKKGQDILLPQM--SGLDERKNACPMHYDFSRQKVSHTTFGSHLCLGQHLAR 365

DB 303 DLDGTVIPAGDTVLVWJADAHRTPEPDPHREDIRRDTAGHLAFGCHGIFCIGAPLAR 362

QY 366 EIVVLKWLTRIPDS--IAPCAQIOHKSIGVSGVQALPLW 406

DB 363 EARIAVALLERCPDIALDVSGELVWYNPMIRGLKAUPIRW 405

RESULT 5

US-09-252-991A-17836

Sequence 17836; Application US/09252991A

Patent No. 6551795

GENERAL INFORMATION:

APPLICANT: Marc C. Rubenfield et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/09/252,991A

PRIOR FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: US 60/074,788

PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190

PRIOR FILING DATE: 1998-07-27

NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 17836

LENGTH: 437

TYPE: PRT

ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-17836

Query Match 16.3%; Score 355; DB 4; Length 437;

Best Local Similarity 25.8%; Pred. No. 4.8e-25;

Matches 96; Conservative 68; Mismatches 158; Indels 50; Gaps 8;

QY 63 WIATRGQIREAYEDYRHFSSCPPIPREA-----GEAYDFIPTSMD 104

DB 60 WYVTR-----YRDARKVLNH--PGVRERQAAELYAKRTGTSFRAGIGEGSHHMLND 111

QY 105 PPEQORFRALANOVGMVVDKLENRIQELACSLIESLRPQGCNFTEDYAEPPPIRIFM 164

DB 112 PDHTLRSLVGAFTRQVERLCQPHRIETELLDMAGREQADLMADFAILTLTAVIF 171

QY 165 LLAGLPEEDIPHLKYTD---QWTRPDGSMTPFAEAKLYDYLIPIIEQRQKPGTDAIS 221

DB 172 ELLGIPEAREHARQSWERQAEILLSPEEQALADAQ---VDYLRVLLEAKRRQPADVYS 228

QY 222 IVANGVNGRPITTSDEAKRMCGLLVGLDVTNVFLSFSMEFLAKSPHROELIERPELI 281

DB 229 GLVQAADSGQSLEAELVSNHLMISGFTTMMIGNALVTLLVNPBQALLRQAQPELL 288

QY 282 PAACEELLRRFSLVADG--RILTSDFEFGVQKKGQDILLPQMLSGIDRKNACPMHVD 339

DB 289 PNAMEELVRHDSVPRASMLRFTVEDVLDGVTTPAGEYILVSNLTANHDAERFDDPDL 348

QY 340 FSRQKVSHTTFGSHLCLGQHLARREIIVTLKWLTRIPDEFSIA--PGAQIOHKSIGVSG 398

DB 349 LTRNTDGHGFGVGVHCVGASLARLEGRITAIQELLARFPDLQIAVPHAEIQ----- 400

QY 399 VQALPLVWDPAT 410

DB 401 -----WLPIT 405

RESULT 6

5212296-9

Patent No. 5212296

APPLICANT: DEAN, CAROLINE; HARDER, PATRICIA A.; LEYO, KENNETH

J.; O'KEEFE, DANIEL P.; OMER, CHARLES A.; ROMESSER, JAMES A.

TEPPERMAN, JAMES M.

TITLE OF INVENTION: EXPRESSION OF HERBICIDE METABOLIZING

CYTOKROMES

NUMBER OF SEQUENCES: 19

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/569,781

FILING DATE: 23-AUG-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 464,499

FILING DATE: 12-JAN-1990

APPLICATION NUMBER: 405,605

FILING DATE: 11-SEP-1989

SEQ ID NO: 9

LENGTH: 403

5212296-9

Query Match 15.9%; Score 346; DB 6; Length 403;

Best Local Similarity 29.2%; Pred. No. 3.9e-28;

Matches 106; Conservative 63; Mismatches 164; Indels 30; Gaps 13;

QY 63 WIATRGQIREAYEDYRHFSSCPPIPREA-----GEAYDFIPTSMDPQORQFRALA 115

DB 52 WLVTIRHQDVRAVLGDER--FSADAHRTGFPFLTAGGRIIIGTNTFTLRMDDEPHARLRML 110

QY 116 NOVGMVVDKLENRIQELACSLIESLRP--QGQCNTEDYAEPPPIRIFMLLAGLPEEDI 174

DB 111 TADFIYKVEAMSPVQRLADLVDRMTTGTSGADIVTEFALPLPLSVICLLIGVPYEDH 170

QY 175 PHL-----KYLTQMTRPDGSMTFAEAKLYDYLIPIIEQRQKPGTDAIS--IVANGVN 229

DB 171 AFFQERSRVLLTLRS--PE---EYRAAQDELLLEYLARLARTKRERPDDAIISRLVARGELD 227

QY 230 GRPITTSDEAKRMCGLLVGLDVTNVFLSFSMEFLAKSPHROELIERPELIPAAACELL 289

DB 228 DTQIAT-----MGRLLLVAGHETTANMTALSTLVLLNPNQALRAEPALVKGAVELL 282

QY 290 RRFSLVADG--RILTSDFEFGVQKKGQDILLPQMLSGIDRKNACP--MHVDFSRQKV 345

DB 283 RYLTIVHNGVPIATIEDVLIGRTIAGEGVLC--MISSANRDAEVPFGDDLDVADAR 340

QY 346 SHTFGSHLCLGQHLARREIIVTLKWLTRIPDEFSIA--PGAQIOHKSIGVSGVQALP 403

DB 341 RHVAFGFGVHCQGLQAPLARVELQIA:ETLLRLRLDLRLAVPHBEEIPRGMAYGVESDP 400

QY 404 LWV 406

DB 401 IAW 403

RESULT 7

US-09-385-028-12

Sequence 12; Application US/09385028

Patent No. 6232106

GENERAL INFORMATION:

APPLICANT: Susan E. Jensen  
APPLICANT: Kwamena A Aidoo  
APPLICANT: Ashish S. Paradkar  
TITLE OF INVENTION: DNA Sequence Encoding Enzymes of Clavulanic  
Patent No. 6232106  
TITLE OF INVENTION: Acid Biosynthesis  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: JACOBSON, PRICE, HOLMAN & STERN, PLLC  
STREET: The Jennifer Building, 400 Seventh Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/385,028  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/790,462  
FILING DATE: 29-JAN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: D. Douglas Price  
REGISTRATION NUMBER: 24,514  
REFERENCE/DOCKET NUMBER: 1418/P57452US2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 638-6666  
TELEFAX: (202) 39305350  
TELEX: RCA 248593 IDEA UR  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 409 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-385-028-12

Query Match 15.6%; Score 339; DB 3; Length 409;  
Best Local Similarity 29.2%; Pred. No. 2.2e-27;  
Matches 123; Conservative 60; Mismatches 184; Indels 54; Gaps 18;  
QY 16 PHVPEHLVFDPMYFNLSAGVQEAHVLQESNVPDLVWTRCNGH-WIATRGQLIREA 74  
DB 14 PAYPMHRVCPVD---PPQLAGLSRQKASRVT-----LW---DGSQVWLVTSHAGARAV 62  
QY 75 YEDYRHS-SECFIP-----REAGAYDFIPTSDPPEQQRQFRA-----LANQV 118  
DB 63 LGRRTTAVTSAGFGFMLTTSQVLVRANPESASFI--RMDDPQHSRLSMLTRDFLARRA 120  
QY 119 VGM-PVVYKLENRIQELACSLIESLRPQGCNFTEDYAEPPPIRIFMLLAGLPEEDIPHL 177  
DB 121 EALRPVAVREL---LDELGLGVKGERP---VDLVAGLTIPVPSRVITLLFGAGDRREFI 174  
QY 178 K-----YLTQMTWRPDGSMTEFAEALYDYLPIIEORRQKPGTDAISIVANGQVGRPI 233  
DB 175 EDRSAVLIDRGYTP---QVAKARDELGVRLVEERIEPNPGTDLISRLWIDQVRPGHL 231  
QY 234 TSEAKEMCGLLVGGDLTVNFLSFMFLAKSPERHQELIERPELIPACELLRFS 293  
DB 232 RVEMVPMCLLLVAGHTTTSQASLSLLTDPGLAGRLTEDPALPKAVELLRFHS 291  
QY 294 LVADG--RIITSDYEEHGVOLKGDQILLPQLSLGLDERKKNACPMYDFSRQKVSHTTFG 351  
DB 292 IVQNGLAARAAVEDVQDDVLIRAGEGVVLSAGNRDETVEPPDVRVDDBDARHLAFG 351  
QY 352 HGHSLCAGQHLAR---REIIVTLKEWLTTRIPDSIA-PGAQIQHKSQIVS-GVOALPLVW 406

DB 352 EEMHQCLGQWLARVELEIILAAVLRW---PGARLAVPPELDFRHEVSSVGLCALPVTW 408  
QY 407 D 407  
DB 409 Z 409  
RESULT 8  
US-09-726-614-12  
Sequence 12, Application US/09726614  
Patent No. 6514735  
GENERAL INFORMATION:  
APPLICANT: Susan E. Jensen  
APPLICANT: Kwamena A Aidoo  
APPLICANT: Ashish S. Paradkar  
TITLE OF INVENTION: DNA Sequence Encoding Enzymes of Clavulanic  
Patent No. 6514735  
TITLE OF INVENTION: Acid Biosynthesis  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: JACOBSON, PRICE, HOLMAN & STERN, PLLC  
STREET: The Jennifer Building, 400 Seventh Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/726,614  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/790,462  
FILING DATE: 29-JAN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: D. Douglas Price  
REGISTRATION NUMBER: 24,514  
REFERENCE/DOCKET NUMBER: 1418/P57452US2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 638-6666  
TELEFAX: (202) 39305350  
TELEX: RCA 248593 IDEA UR  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 409 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-726-614-12

Query Match 15.6%; Score 339; DB 4; Length 409;  
Best Local Similarity 29.2%; Pred. No. 2.2e-27;  
Matches 123; Conservative 60; Mismatches 184; Indels 54; Gaps 18;  
QY 16 PHVPEHLVFDPMYFNLSAGVQEAHVLQESNVPDLVWTRCNGH-WIATRGQLIREA 74  
DB 14 PAYPMHRVCPVD---PPQLAGLSRQKASRVT-----LW---DGSQVWLVTSHAGARAV 62  
QY 75 YEDYRHS-SECFIP-----REAGAYDFIPTSDPPEQQRQFRA-----LANQV 118  
DB 63 LGRRTTAVTSAGFGFMLTTSQVLVRANPESASFI--RMDDPQHSRLSMLTRDFLARRA 120  
QY 119 VGM-PVVYKLENRIQELACSLIESLRPQGCNFTEDYAEPPPIRIFMLLAGLPEEDIPHL 177  
DB 121 EALRPVAVREL---LDELGLGVKGERP---VDLVAGLTIPVPSRVITLLFGAGDRREFI 174  
QY 178 K-----YLTQMTWRPDGSMTEFAEALYDYLPIIEORRQKPGTDAISIVANGQVGRPI 233

Db 175 EDSAVLIDRGTYPE---QVAKARDELGDYLRNLEBBERIENPGTDLISRLVLDQVRPGLH 231  
QY 234 TSDAKRMCGLLVGGDLTVNLFSPMEPLAKSPHROELIERPELIPACBELLARFS 293  
Db 232 RVEEMVPMCRLLVAGHGTTSQASLSLLTDPPELAGRLTEDPALLPKAVBELLARFS 291  
QY 294 LVADG--RLITSDFEHGVOLKKGDOILLPQMLSGDLDERKNACPMHVDPSRQKVSHTTFG 351  
Db 292 IVQNGLARAAVEDVQLDVLIRAGEGVVLSLSAGNRDETVPDPDRVDVDRDARHLAFG 351  
QY 352 HGSHLCGLQGHAR---REIIVT-KENLIRIPDFSIA-PCAQIQHKSGIVS-GVQALPLVW 406  
Db 352 HGMHCQCLQWLARVELEELAAVLRW---PGARLAVPFEELDFRHEVSSVGLGALPVTW 408  
QY 407 D 407  
Db 409 Z 409

RESULT 9  
US-09-385-040-12  
; Sequence 12, Application US/09385040  
; Patent No. 6589775  
; GENERAL INFORMATION:  
; APPLICANT: Jensen, Susan E  
; APPLICANT: Aidoo, Kwamena A  
; APPLICANT: Paradkar, Ashish S  
; TITLE OF INVENTION: DNA SEQUENCE ENCODING ENZYMES OF CLAVULANIC ACID  
; FILE REFERENCE: 09/385,040  
; CURRENT FILING DATE: 1999-08-30  
; PRIOR APPLICATION NUMBER: US/09/385,040  
; PRIOR FILING DATE: 1997-01-29  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 12  
; LENGTH: 409  
; TYPE: PRT  
; ORGANISM: Streptomyces clavuligerus  
US-09-385-040-12

Query Match 15.6%; Score 339; DB 4; Length 409;  
Best Local Similarity 23.2%; Pred. No. 2.2e-27;  
Matches 123; Conservative 60; Mismatches 184; Indels 54; Gaps 18;

QY 16 PHVPEHLVDFDYMNPNSLAVQOAVLQESNVPLDVWTRCNGH-WIATRGQLIREAYEDYR 74  
Db 14 PAVPMHRVCPVD---PPQLAGLSRQKAASRVT---LW---DGSQVWLVTSHAGARAV 62  
QY 75 YEDYRHS-SECPFIP-----REAGEAYDPIPTSMOPPEQOEFA-----LANQV 118  
Db 63 LGDRRFATVTSAPGFPMLTRTSQVLRANPESAFI--RMDFQHSRLRSM/CRDFLARRA 120  
QY 119 VGM--PVYDKLENRIQELACSLIESLRPGQCNFTEDYAEPPPIRIFMLLAGLPEEDIPHL 177  
Db 121 EALRPVAREL---LDELGLGVKSERP---VDLVAGLTIPVESRVTLLFGAGDREFFI 174  
QY 178 K-----YLTQMTRPDGSMTFAEAKALVDYLIPIIEQRORKEPTDAISIVANGQVNGRPI 233  
Db 175 EDRSAVLIDRGTYPE---QVAKARDELGDYLRNLEBBERIENPGTDLISRLVLDQVRPGLH 231  
QY 234 TSDAKRMCGLLVGGDLTVNLFSPMEPLAKSPHROELIERPELIPACBELLARFS 293  
Db 232 RVEEMVPMCRLLVAGHGTTSQASLSLLTDPPELAGRLTEDPALLPKAVBELLARFS 291  
QY 294 LVADG--RLITSDFEHGVOLKKGDOILLPQMLSGDLDERKNACPMHVDPSRQKVSHTTFG 351  
Db 292 IVQNGLARAAVEDVQLDVLIRAGEGVVLSLSAGNRDETVPDPDRVDVDRDARHLAFG 351  
QY 352 HGSHLCGLQGHAR---REIIVT-KENLIRIPDFSIA-PCAQIQHKSGIVS-GVQALPLVW 406  
Db 352 HGMHCQCLQWLARVELEELAAVLRW---PGARLAVPFEELDFRHEVSSVGLGALPVTW 408

QY 407 D 407  
Db 409 Z 409

RESULT 10  
5212296-6  
; Patent No. 5212296  
; APPLICANT: DEAN, CAROLINE HARDER, PATRICIA A. LETO, KENNETH  
; J.; O'KEEFE, DANIEL P.; OMES, CHARLES A.; ROMESSER, JAMES A.  
; TEPPEMAN, JAMES M.  
; TITLE OF INVENTION: EXPRESSION OF HERBICIDE METABOLIZING  
; CYTOCHROMES  
; NUMBER OF SEQUENCES: 19  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/569,781  
; FILING DATE: 23-AUG-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 464,499  
; FILING DATE: 12-JAN-1990  
; APPLICATION NUMBER: 405,605  
; FILING DATE: 11-SEP-1989  
; SEQ ID NO: 6  
; LENGTH: 406  
5212296-6

Query Match 15.4%; Score 336; DB 6; Length 406;  
Best Local Similarity 26.6%; Pred. No. 4.6e-27;  
Matches 105; Conservative 74; Mismatches 192; Indels 24; Gaps 11;

QY 31 PSNLSAGVQ--EAWAVLQESNVPLDVWTRCNGH-WIATRGQLIREAYEDYR----- 79  
Db 17 PSNRSCPQLDPDGYAQLRDTGPHLRVTVLYDGRQAWVTKEAMARKLGLDPLSSNRDTD 76  
QY 80 HFSSECPFIP--REAGEAYDPIPTSMOPPEQOEFAALANQVVMVVDKLENRIQELACS 137  
Db 77 NFPATSPRFAEVRESPOAF---IGLDPPEHGTTRMTISEFTVKRIKMRPEVEVHVHG 132  
QY 138 LIESLRPOS-OCNFTEDYAEPPPIRIFMLLAGLPEEDIPHLKYLTDQMTRPDGSMTFAEA 196  
Db 133 FLDEMLAAGPTADLVSQFALPVPMSVICRLLGVPYADHEFFDQASKRLVQSTDAQSALTA 192  
QY 197 KEAYDYLIPIIEQRORKEPTDAI-SIVANGQVNGRPIITSDAEKRMCGLLVAGGLDVTWN 255  
Db 193 RNDLAGYLDGLITQFQTEFGAGLVGALVADQLANGE-IDRSELISTAMLLAGHETTAS 251  
QY 256 FLTSMEFLAKSPHROELIERPELIPACBELLARFSL--VADGRILTSDFEHGVOLK 313  
Db 252 MTSLSVITLLDHPCEYAAALRADRSILVPGAVEVELLYALADIAGGRVATADIEVEGHLR 311  
QY 314 KGDQILLPQMLSGDLDERKNACPMHVDPSRQKVSHTTFGHSGLCLGOLARREIIVTKE 373  
Db 312 AGEVIVVNSIANRDTGVYEDPDALDIHRSARHHLAFGFGVHCQCLQNLARLELEVLNA 371  
QY 374 WLTRIPDFSIA-PCAQIQHKSG-IVSGVQALPLVW 406  
Db 372 LMDRVTPLRLAVPVEQLVLRPGTIIQSVNELPVTW 406

RESULT 11  
US-08-765-907A-10  
; Sequence 10, Application US/08765907A  
; Patent No. 6352839  
; GENERAL INFORMATION:  
; APPLICANT: BLANC, Veronique  
; APPLICANT: THIBAUT, Denis  
; APPLICANT: BAMAS-JACQUES, Nathalie  
; APPLICANT: BLANCHE, Francis  
; APPLICANT: COUZET, Joel  
; APPLICANT: BARRIERE, Jean-Claude  
; APPLICANT: DEBUSSCHE, Laurent  
; APPLICANT: FAMECHON, Alain

```

; APPLICANT: PARIS, Jean-Marc
; APPLICANT: ETRUC-ROSSET, Gilles
; TITLE OF INVENTION: Streptogramins And Method For Preparing Same By
; TITLE OF INVENTION: Mutasynthesis
; FILE REFERENCE: Streptogramin genes
; CURRENT APPLICATION NUMBER: US/08/765,907A
; CURRENT FILING DATE: 1997-03-20
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 10
; TYPE: PRT
; ORGANISM: Streptomyces pristinaespiralis
US-08-765-907A-10

Query Match: 14.8%; Score 323.5; DB 4; Length 399;
Best Local Similarity 29.0%; Pred. No. 9.6e-26;
Matches 106; Conservative 59; Mismatches 159; Indels 41; Gaps 12;

QY 74 AYEDYRH-----FSSECFIPREAGEAYDRIPTSMPPQORQFRALANQVGM 121
DB 36 AHWFRADVLTASDPGVSSQSLRPPGQALSEQILSVDPDMHRTLRVLSQAFTP 95

122 PVDKLENRQELACSLIESLRPOQC-NFTEDYAEFPPIRIFMLLAGLPBEDIPHLKYL 180
DB 96 RTVADLEPRVTELAGQLDAV--DGDIFLIWADFAYPPLVIVIAELLGVPPADRTLFRSW 153

181 TDQMT-----PGSMTFAZAKALYDYLPIIPIEQRKQGTDA-SIVA 224
DB 154 SDRMLQWQADPADMQFGDDADEYQRLVKEPFRAMHAYLHDVTDRRARPANL-SALV 213

225 NGQVNGRPITSDEAKRMCGLLVGLDVTWNFLSFMFLAKSPEHRQELIERPE--LIP 282
DB 214 AARVEGRUTDEQIVERGALCIMAGHVSMTMLGNVTULCKDHP--RAEAAARADRSLLP 271

283 AACBELLR-RFLSLVADGRILTSDEYFHGVOLKKGQDILLPQVLS-GLDERKNACPMHVD 340
DB 272 ALIEEVLRLPPIITWAVTTKTDTVLAGTTIPAG-RMVVPSLLSANHDEQVETDPDHLDL 330

341 SRQKVSHTTFHGSHLCLGQHLARREIIVTLKWLTRIPDSIAPGAIQ-HKSGIYSGV 399
DB 331 AREG-RQIAFGHGHYCLGASLARLEGRALAEALFDRFPDSPTDGAKLRYHRDGLF-GV 388

400 QALPL 404
DB 389 KNLPL 393

```

```

RESULT 12
US-08-102-863-11
; Sequence 11, Application US/08102863
; Patent No. 5466590
; GENERAL INFORMATION:
; APPLICANT: SARIASLANI, S.M.A
; TITLE OF INVENTION: CONSTITUTIVE
; TITLE OF INVENTION: EXPRESSION OF P450SOY
; TITLE OF INVENTION: AND FERREDOXIN-SCY IN
; TITLE OF INVENTION: STREPTOMYCES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS
; ADDRESS: AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: USA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0,
; SOFTWARE: Version #1.25

```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/102,863
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/807,001
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: GALLEGOS, R. THOMAS
; REGISTRATION NUMBER: 32,692
; REFERENCE/DOCKET NUMBER: CR-9000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-892-7342
; TELEFAX: 302-892-7949
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 412 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-102-863-11

Query Match 14.5%; Score 315; DB 1; Length 412;
Best Local Similarity 27.2%; Pred. No. 8.2e-25;
Matches 84; Conservative 56; Mismatches 163; Indels 6; Gaps 5;

QY 103 MDPPEQQRALANQVGMPPVVDKLENRQELACSLIESLRPOG-QCNFTEDYAEFPPIR 161
DB 105 VDDPEHMTQRRLMPTTFSKXIGALRPRIQETVVDLDMERQGPALVSAFALPVFSM 164

162 IFMLLAGLPBEDIPHLKYLTDQMTRPDGSMTFAEAKALYDYLPIIPIEQRKQGTDAIS 221
DB 165 VICALLGVYADHAFFERSQRLLRGCGADDVNRARDEEYVIGALIDRKRAEPDGLLD 224

222 IVANGQVNGRPITSDEAKRMCGLLVGLDVTWNFLSFMFLAKSPEHRQELIERPELI 281
DB 225 ELIHRDHPDGVDRQELVAFVILLIAGHETTANNISLGTFTLLSHBPQLAALRAGGTST 284

282 PAACEELLRRFLSLVADG--RLTSDYEPHGVOLKKGQDILLPQVLSGLDERKNACPMHVD 339
DB 285 AVVVEELL-RFLSLAEGRLATDEMEVDGATIRKGEVGFSTSLNRDADVFPRAETLD 343

340 FSROKVSHTTFHGSHLCLGQHLARREIIVTLKWLTRIPDPSIA-PGAIQHKSG-IVS 397
DB 344 WDRPARHHLAFGFGVHQCLQGLMARAEIDIARTLFRPLGLRLAVPAHEIRHKPGDTIQ 403

398 GVQALPLVW 406
DB 404 GLLDLPVAV 412

```

```

RESULT 13
PCT-US92-10885-11
; Sequence 11, Application PC/TUS9210885
; GENERAL INFORMATION:
; APPLICANT: SARIASLANI, SIMA
; TITLE OF INVENTION: CONSTITUTIVE
; TITLE OF INVENTION: EXPRESSION OF P450SOY
; TITLE OF INVENTION: AND FERREDOXIN-SOY IN
; TITLE OF INVENTION: STREPTOMYCES
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS
; ADDRESS: AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: USA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch,
; MEDIUM TYPE: 1.0 MB

```

```

;
; COMPUTER: Macintosh
; OPERATING SYSTEM: Macintosh System, 6.0
; SOFTWARE: Microsoft Word, 4.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/10885
; FILING DATE: 199212-6
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: GALLEGOS, R. THOMAS
; REGISTRATION NUMBER: 32,692
; REFERENCE/DOCKET NUMBER: CR-9000-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-892-7342
; TELEFAX: 302-892-7949
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 412 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; PCT-US92-10885-11

Query Match      14.5%; Score 315; DB 5; Length 412;
Best Local Similarity 27.2%; Pred. No. 8.2e-25;
Matches 84; Conservative 56; Mismatches 163; Indels 6; Gaps 5;

QY 103 MDPPEQGFRLANQVGMVVDKLENRIQELACSLIESLRPQG-OCNFTEDYAEPPPIR 161
DQ 105 VDDPHNTQRMPLTPTFSKRGIGALRPRIQETVDRLDDAMERQGPFAELVAFALPVPF 164
QY 162 IFMLLAGLPEDIPHLKYLCTQMTPEPDGSMTFAEAKYALDYLIPIIQRQKQCTDAIS 221
DQ 165 VICALLGVYADHAFPEERSQRLRGDGVNRRARDEEYLGALIDRKRAEPDGLD 224
QY 222 IVANGVNGRPITTSDEAKRMGLLVGGGLTVVNFSLFSMEFLAKSPHRCQLIERPELI 281
DQ 225 ELIERHDHDPGVDRQLVAFVILLIAGHETTANNISLGLTFLLSHPQLAALRAGGTST 284
QY 282 PAACEELLRRSLVADG--RLTSDYEFHGVOLKGGDQILLPQMLSGLDERKNACPMHVD 339
DQ 285 AVVVEELL-RLSLAEGQLRLATEDMEVDGATIRKGGVVFSTSLINRDADVFPRAETLD 343
QY 340 FSRQKVSHTTGCHSHLCLGHLARPELIIVTLKWLTRIPDSIA-PGNAQLOKSG-IVS 397
DQ 344 WDRPARHHLAFGFGVHQCLGQKLARAEIDIAKRTLFERLPGLRLAVPAHEIRHKPGDTIQ 403
QY 398 GWQALPLVW 406
DQ 404 GLLDLFVAV 412

RESULT 14
US-09-335-409-8
; Sequence 8, Application US/09335409
; Patent No. 6121029
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/335,409
; CURRENT FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 8
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Sorangium cellulosum

Query Match      12.7%; Score 277.5; DB 3; Length 419;
Best Local Similarity 24.0%; Pred. No. 8.4e-21;
Matches 101; Conservative 67; Mismatches 178; Indels 75; Gaps 14;

US-09-335-409-8
Query Match      12.7%; Score 277.5; DB 3; Length 419;
Best Local Similarity 24.0%; Pred. No. 8.4e-21;
Matches 101; Conservative 67; Mismatches 178; Indels 75; Gaps 14;

QY 6 OSNANLAPLPHVPEHLVDFDMYNPSNLISAGVQE-----AWAVLOESNVVDLWTRCNEG 61
DQ 3 QEQANQSETKP-----AFDFKFPAP-----GYADFFPAIERLREA-TPIFYWD--EGR 48
QY 62 HWIATRGQ-----LIREAYDYRHFSSCPPIPREAGAYDTPTSMOPPEQR 109
DQ 49 SWULTRYHDVSAVFRDERFAVSKEEWSAEYSSAIP-----ELSDMKKXGLFGLPPEDHA 104
QY 110 QERALANQVGMVVDKLENRIQELACSLIESLRPQGCNFTEDYAEPEIRIFMLAGL 163
DQ 105 KVRKLVNPSFTSRAIDLRLRAEIQRTVDQLLDARSQGEEDFVVRDYAEIGPRAISALLKV 164
QY 170 PEEDIPHLKYLTDQMTDPDGSMTFAEAKYALDYLIPI-----I 207
DQ 165 PAE-----CDEKERRFGSAT-----ARALGVGLVPQVDDEETKTLVASVTGCLALLHDV 212
QY 208 IEQREQKP-GTDAISIVANGVNGRPITTSDEAKRMGLLVGGGLTVVNFSLFSMEFLAK 266
DQ 213 LDERRRNPLENDVLTMLQAEADGSRSLTKELVALVGAIIAAGTDTTIVLIAFAVLNLR 272
QY 267 SPEHROQLIERELIPAAACEELLRRFSLVADG--RLTSDYEFHGVOLKGGDOI--LLPQ 322
DQ 273 SPEALVKAEPLMRNALDEVLRFDNLIRIGTFARQDLEYCGASIKKGENVFLIPS 332
QY 323 MLSGGLDERKNACPMHVDVFSRQKVSHTTGHGSHLCLGHLARPELIIVTLKWLTRIPDS 382
DQ 333 ALR--CGTVFSPDPVDVDRDTCASLAYGRGHVCPGVSLARLEAELANGTIFERRPEMK 390
QY 383 I 383
DQ 391 L 391

RESULT 15
US-09-413-814-71
; Sequence 71, Application US/09413814
; Patent No. 6225064
; GENERAL INFORMATION:
; APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH
; APPLICANT: Bristol-Myers Squibb, Co.
; APPLICANT: Beyer, Stefan
; APPLICANT: Bloecker, Helmut
; APPLICANT: Brandt, Petra
; APPLICANT: Cino, Paul M
; APPLICANT: Dougherty, Brian A
; APPLICANT: Goldberg, Steven L
; APPLICANT: Hofle, Gerhard
; APPLICANT: Mueller, Joachim
; APPLICANT: Reichenbach, Hans
; TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polyketide or
; FILE REFERENCE: PCT/US 99/23535
; CURRENT APPLICATION NUMBER: US/09/413,814
; CURRENT FILING DATE: 1999-10-07
; EARLIER APPLICATION NUMBER: DE 198 46 493.2
; EARLIER FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 71
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Sorangium cellulosum

US-09-413-814-71
Query Match      12.7%; Score 277.5; DB 3; Length 419;
Best Local Similarity 24.0%; Pred. No. 8.4e-21;
Matches 101; Conservative 67; Mismatches 178; Indels 75; Gaps 14;
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 6, 2004, 18:53:57 ; Search time 34.974 Seconds

(without alignments)  
3108.863 Million cell updates/sec

Title: US-09-246-451A-13

Perfect score: 2177

Sequence: 1 TTTTQSNANLAPLPHVFE.....IVSGUQALPLVNDPATKAV 414

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1071772 seqs, 262633353 residues

Total number of hits satisfying chosen parameters: 1071772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA: \*

1: /cgn2\_6/prodata/1/pubpaa/US07\_PUBCOMB.pep: \*  
2: /cgn2\_6/prodata/1/pubpaa/PCT\_NEW\_PUB.pep: \*  
3: /cgn2\_6/prodata/1/pubpaa/US06\_NEW\_PUB.pep: \*  
4: /cgn2\_6/prodata/1/pubpaa/US06\_PUBCOMB.pep: \*  
5: /cgn2\_6/prodata/1/pubpaa/US07\_NEW\_PUB.pep: \*  
6: /cgn2\_6/prodata/1/pubpaa/PCTUS\_PUBCOMB.pep: \*  
7: /cgn2\_6/prodata/1/pubpaa/US08\_NEW\_PUB.pep: \*  
8: /cgn2\_6/prodata/1/pubpaa/US08\_PUBCOMB.pep: \*  
9: /cgn2\_6/prodata/1/pubpaa/US09A\_PUBCOMB.pep: \*  
10: /cgn2\_6/prodata/1/pubpaa/US09B\_PUBCOMB.pep: \*  
11: /cgn2\_6/prodata/1/pubpaa/US09C\_PUBCOMB.pep: \*  
12: /cgn2\_6/prodata/1/pubpaa/US09\_NEW\_PUB.pep: \*  
13: /cgn2\_6/prodata/1/pubpaa/US10A\_PUBCOMB.pep: \*  
14: /cgn2\_6/prodata/1/pubpaa/US10B\_PUBCOMB.pep: \*  
15: /cgn2\_6/prodata/1/pubpaa/US10C\_PUBCOMB.pep: \*  
16: /cgn2\_6/prodata/1/pubpaa/US10\_NEW\_PUB.pep: \*  
17: /cgn2\_6/prodata/1/pubpaa/US60\_NEW\_PUB.pep: \*  
18: /cgn2\_6/prodata/1/pubpaa/US60\_PUBCOMB.pep: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2177	100.0	414	15	US-10-453-104-13
2	2169	99.6	414	15	US-10-453-104-11
3	2165	99.4	414	15	US-10-453-104-2
4	2162	99.3	414	15	US-10-453-104-12
5	372	17.1	399	14	US-10-314-657-37
6	366.5	16.8	404	14	US-10-214-446-50
7	365.5	16.8	404	14	US-10-214-446-40
8	365	16.8	416	14	US-10-156-761-14997
9	363.5	16.7	416	9	US-09-861-289-39
10	363.5	16.7	416	10	US-09-860-846-39
11	363.5	16.7	416	10	US-09-388-3848-39
12	363.5	16.7	416	10	US-09-836-821-39
13	363.5	16.7	416	10	US-09-793-778-18
14	363.5	16.7	416	14	US-10-201-365-13
15	363.5	16.7	416	14	US-10-160-539-18

Sequence 33, Appl  
Sequence 38, Appl  
Sequence 9525, Ap  
Sequence 20, Appl  
Sequence 9914, Ap  
Sequence 8376, Ap  
Sequence 559, Ap  
Sequence 11073, A  
Sequence 13776, A  
Sequence 4, Appl  
Sequence 3, Appl  
Sequence 2, Appl  
Sequence 8710, Ap  
Sequence 9703, Ap  
Sequence 6, Appl  
Sequence 12, Appl  
Sequence 15, Appl  
Sequence 22, Appl  
Sequence 14, Appl  
Sequence 14659, A  
Sequence 18, Appl  
Sequence 35, Appl  
Sequence 18431, A  
Sequence 46, Appl  
Sequence 22, Appl  
Sequence 8, Appl  
Sequence 6, Appl  
Sequence 4117, Ap  
Sequence 7959, Ap  
Sequence 14, Appl

US-10-271-889-39  
US-10-214-446-38  
US-10-156-761-9525  
US-10-214-446-20  
US-10-156-761-9914  
US-10-156-761-8376  
US-10-389-647-559  
US-10-156-761-11073  
US-10-156-761-13776  
US-10-214-446-4  
US-10-214-446-32  
US-10-214-446-2  
US-10-156-761-8710  
US-10-156-761-9703  
US-10-201-213-6  
US-10-458-201-12  
US-10-214-446-16  
US-10-214-446-22  
US-10-214-446-14  
US-10-156-761-14659  
US-10-214-446-18  
US-10-156-761-10431  
US-10-214-446-46  
US-10-145-415-22  
US-10-335-032-8  
US-10-145-415-6  
US-09-738-826-4117  
US-10-156-761-7959  
US-10-145-415-14

#### ALIGNMENTS

#### RESULT 1

US-10-453-104-13  
; Sequence 13, Application US/10453104  
; Publication No. US20030207345A1  
; GENERAL INFORMATION:  
; APPLICANT: California Institute of Technology;  
; APPLICANT: Frances H. Arnold  
; APPLICANT: Hyun Joo  
; TITLE OF INVENTION: Oxygenase Enzymes and Screening Method  
; FILE REFERENCE: 4058/1827-US3  
; CURRENT APPLICATION NUMBER: US/10/453,104  
; PRIOR FILING DATE: 2003-06-02  
; PRIOR APPLICATION NUMBER: US 09/661,093  
; PRIOR FILING DATE: 2000-09-13  
; PRIOR APPLICATION NUMBER: US 09/246,451  
; PRIOR FILING DATE: 1999-02-09  
; PRIOR APPLICATION NUMBER: US 60/094,403  
; PRIOR FILING DATE: 1998-07-28  
; PRIOR APPLICATION NUMBER: US 60/106,840  
; PRIOR FILING DATE: 1998-11-03  
; PRIOR APPLICATION NUMBER: US 60/086,206  
; PRIOR FILING DATE: 1998-05-21  
; PRIOR APPLICATION NUMBER: US 60/106,834  
; PRIOR FILING DATE: 1998-11-03  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: FASTSEQ for Windows Version 3.0  
; SEQ ID NO 13  
; LENGTH: 414  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Mutant M7-8H  
US-10-453-104-13

Query Match 100.0%; Score 2177; DB 15; Length 414;

Best local Similarity 100.0%; Pred. NC. 9.8e-211;

Matches 414; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



QY 1 TTETIQSNANLAPLPPHVPHELVDFDQYXNPSNLSAGVQEAQAVLQESNVDPDLVWTRCNG 60  
DB 1 TTETIQSNANLAPLPPHVPHELVDFDQYXNPSNLSAGVQEAQAVLQESNVDPDLVWTRCNG 60  
QY 61 GHWIATRGQLIREAYEDYRHSFSECPFIIPREAGEAYDFIPTSMPPPEQORFRALANQVVG 120  
DB 61 GHWIATRGQLIREAYEDYRHSFSECPFIIPREAGEAYDFIPTSMPPPEQORFRALANQVVG 120  
QY 121 MPVVDKLENRIQELACSLIESLRPOQCNTFYAEPPPIRIFMLLAGLPEEDIPHLKYL 180  
DB 121 MPVVDKLENRIQELACSLIESLRPOQCNTFYAEPPPIRIFMLLAGLPEEDIPHLKYL 180  
QY 181 TDQMTRPDGSMTFAEAKAALYDYLPIIEORROKPGTDAISIVANGQVNGRPITSDEAKR 240  
DB 181 TDQMTRPDGSMTFAEAKAALYDYLPIIEORROKPGTDAISIVANGQVNGRPITSDEAKR 240  
QY 241 MFGLLLVGGGLDVTNVFLSFSMEFLAKSPEHRQELIERPERIPAAACELLRRFSLVADGRI 300  
DB 241 MFGLLLVGGGLDVTNVFLSFSMEFLAKSPEHRQELIERPERIPAAACELLRRFSLVADGRI 300  
QY 301 LTSDFYFHGVQLKKGQOILLPQMLSGLDERKNACPMHVDFSROKVSHTTFHGSHLCLGQ 360  
DB 301 LTSDFYFHGVQLKKGQOILLPQMLSGLDERKNACPMHVDFSROKVSHTTFHGSHLCLGQ 360  
QY 361 HLARREIIVTKELWLRIPDPSIAPGAQIOHKSGIVSGVQALPLVWDPATTKAV 414  
DB 361 HLARREIIVTKELWLRIPDPSIAPGAQIOHKSGIVSGVQALPLVWDPATTKAV 414

RESULT 2  
US-10-453-104-11  
; Sequence 11, Application JS/10453104  
; Publication No. US20030207345A1  
; GENERAL INFORMATION:  
; APPLICANT: California Institute of Technology;  
; APPLICANT: Frances H. Arnold  
; APPLICANT: Hyun Joo  
; TITLE OF INVENTION: Oxygenase Enzymes and Screening Method  
; FILE REFERENCE: 4058/1E827-US3  
; CURRENT APPLICATION NUMBER: US/10/453,104  
; CURRENT FILING DATE: 2003-06-02  
; PRIOR FILING DATE: 2000-09-13  
; PRIOR FILING DATE: 1999-02-09  
; PRIOR FILING DATE: 1999-02-09  
; PRIOR FILING DATE: 1998-07-28  
; PRIOR FILING DATE: 1998-07-28  
; PRIOR FILING DATE: 1998-11-03  
; PRIOR FILING DATE: 1998-05-21  
; PRIOR FILING DATE: 1998-11-03  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 11  
; LENGTH: 414  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Mutant M7-4H  
US-10-453-104-11

Query Match 99.6%; Score 2169; DB 15; Length 414;  
Best Local Similarity 99.8%; Pred. No. 6.3e-210;  
Matches 413; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 TTETIQSNANLAPLPPHVPHELVDFDQYXNPSNLSAGVQEAQAVLQESNVDPDLVWTRCNG 60  
DB 1 TTETIQSNANLAPLPPHVPHELVDFDQYXNPSNLSAGVQEAQAVLQESNVDPDLVWTRCNG 60  
QY 61 GHWIATRGQLIREAYEDYRHSFSECPFIIPREAGEAYDFIPTSMPPPEQORFRALANQVVG 120

DB 61 GHWIATRGQLIREAYEDYRHSFSECPFIIPREAGEAYDFIPTSMPPPEQORFRALANQVVG 120  
QY 121 MPVVDKLENRIQELACSLIESLRPOQCNTFYAEPPPIRIFMLLAGLPEEDIPHLKYL 180  
DB 121 MPVVDKLENRIQELACSLIESLRPOQCNTFYAEPPPIRIFMLLAGLPEEDIPHLKYL 180  
QY 181 TDQMTRPDGSMTFAEAKAALYDYLPIIEORROKPGTDAISIVANGQVNGRPITSDEAKR 240  
DB 181 TDQMTRPDGSMTFAEAKAALYDYLPIIEORROKPGTDAISIVANGQVNGRPITSDEAKR 240  
QY 241 MFGLLLVGGGLDVTNVFLSFSMEFLAKSPEHRQELIERPERIPAAACELLRRFSLVADGRI 300  
DB 241 MFGLLLVGGGLDVTNVFLSFSMEFLAKSPEHRQELIERPERIPAAACELLRRFSLVADGRI 300  
QY 301 LTSDFYFHGVQLKKGQOILLPQMLSGLDERKNACPMHVDFSROKVSHTTFHGSHLCLGQ 360  
DB 301 LTSDFYFHGVQLKKGQOILLPQMLSGLDERKNACPMHVDFSROKVSHTTFHGSHLCLGQ 360  
QY 361 HLARREIIVTKELWLRIPDPSIAPGAQIOHKSGIVSGVQALPLVWDPATTKAV 414  
DB 361 HLARREIIVTKELWLRIPDPSIAPGAQIOHKSGIVSGVQALPLVWDPATTKAV 414

RESULT 3  
US-10-453-104-2  
; Sequence 2, Application US/10453104  
; Publication No. US20030207345A1  
; GENERAL INFORMATION:  
; APPLICANT: California Institute of Technology;  
; APPLICANT: Frances H. Arnold  
; APPLICANT: Hyun Joo  
; TITLE OF INVENTION: Oxygenase Enzymes and Screening Method  
; FILE REFERENCE: 4058/1E827-US3  
; CURRENT APPLICATION NUMBER: US/10/453,104  
; CURRENT FILING DATE: 2003-06-02  
; PRIOR FILING DATE: 2000-09-13  
; PRIOR FILING DATE: 1999-02-09  
; PRIOR FILING DATE: 1999-02-09  
; PRIOR FILING DATE: 1998-07-28  
; PRIOR FILING DATE: 1998-07-28  
; PRIOR FILING DATE: 1998-11-03  
; PRIOR FILING DATE: 1998-05-21  
; PRIOR FILING DATE: 1998-11-03  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 414  
; TYPE: PRT  
; ORGANISM: P. Putida  
US-10-453-104-2

Query Match 99.4%; Score 2165; DB 15; Length 414;  
Best Local Similarity 99.5%; Pred. No. 1.6e-209;  
Matches 412; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 TTETIQSNANLAPLPPHVPHELVDFDQYXNPSNLSAGVQEAQAVLQESNVDPDLVWTRCNG 60  
DB 1 TTETIQSNANLAPLPPHVPHELVDFDQYXNPSNLSAGVQEAQAVLQESNVDPDLVWTRCNG 60  
QY 61 GHWIATRGQLIREAYEDYRHSFSECPFIIPREAGEAYDFIPTSMPPPEQORFRALANQVVG 120  
DB 61 GHWIATRGQLIREAYEDYRHSFSECPFIIPREAGEAYDFIPTSMPPPEQORFRALANQVVG 120  
QY 121 MPVVDKLENRIQELACSLIESLRPOQCNTFYAEPPPIRIFMLLAGLPEEDIPHLKYL 180  
DB 121 MPVVDKLENRIQELACSLIESLRPOQCNTFYAEPPPIRIFMLLAGLPEEDIPHLKYL 180  
QY 181 TDQMTRPDGSMTFAEAKAALYDYLPIIEORROKPGTDAISIVANGQVNGRPITSDEAKR 240

Db 181 TDQMPDGSMTFAEAKAALDYLIPIIEQRQKPGTDALSIIVANGQVNGRPITSDEAKR 240  
QY 241 MFCLLVGLDVTVNF-SFSMEFLAKSPEHRQELIERPERIPACBELLRRFSLVADGRI 300  
Db 241 MFCLLVGLDVTVNF-SFSMEFLAKSPEHRQELIERPERIPACBELLRRFSLVADGRI 300  
QY 301 LTSDEYEHGVLKKGQDILLPQMLSGLDERKNACPMHVDTSRQKVSHTTTFGHSHLCLGQ 360  
Db 301 LTSDEYEHGVLKKGQDILLPQMLSGLDERKNACPMHVDTSRQKVSHTTTFGHSHLCLGQ 360  
QY 361 HLAARREIIVTLKEMWLRIPDPSIAPGAQIOHKSGIVSGVQALPLVMDPATTKAV 414  
Db 361 HLAARREIIVTLKEMWLRIPDPSIAPGAQIOHKSGIVSGVQALPLVMDPATTKAV 414

## RESULT 4

US-10-453-104-12  
; Sequence 12, Application US/10453104  
; Publication No. US20030207345A1  
; GENERAL INFORMATION:  
; APPLICANT: California Institute of Technology;  
; APPLICANT: Frances H. Arnold  
; APPLICANT: Hyun Joo  
; TITLE OF INVENTION: Oxygenase Enzymes and Screening Method  
; FILE REFERENCE: 4058/1827-US3  
; CURRENT APPLICATION NUMBER: US/10/453,104  
; CURRENT FILING DATE: 2003-06-02  
; PRIOR APPLICATION NUMBER: US 09/661,093  
; PRIOR FILING DATE: 2000-09-13  
; PRIOR APPLICATION NUMBER: US 09/246,451  
; PRIOR FILING DATE: 1999-02-09  
; PRIOR APPLICATION NUMBER: US 60/094,403  
; PRIOR FILING DATE: 1998-07-28  
; PRIOR APPLICATION NUMBER: US 60/106,840  
; PRIOR FILING DATE: 1998-11-03  
; PRIOR APPLICATION NUMBER: US 60/086,206  
; PRIOR FILING DATE: 1998-05-21  
; PRIOR APPLICATION NUMBER: US 60/106,834  
; PRIOR FILING DATE: 1998-11-03  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: Fast-SEQ for Windows Version 3.0  
; SEQ ID NO 12  
; LENGTH: 4-4  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Mutant M7-6H  
US-10-453-104-12

Query Match 99.3%; Score 2162; DB 15; Length 414;  
Best Local Similarity 99.5%; Pred. No. 3.2e-209;  
Matches 412; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 TLETIQSNANLAPLPHPVHELVDYFDMYFNSLSAGVQEAWAYLQSNVDPDLVWTRCNG 60  
Db 1 TLETIQSNANLAPLPHPVHELVDYFDMYFNSLSAGVQEAWAYLQSNVDPDLVWTRCNG 60  
QY 61 GHWIATRGQIREAYEDYRHFSSCPPIPRAGAYDFIPTSDMPPEQORFRALANQVVG 120  
Db 61 GHWIATRGQIREAYEDYRHFSSCPPIPRAGAYDFIPTSDMPPEQORFRALANQVVG 120  
QY 121 YEVVDKLENRIQELACSLIESLRQGCNFTEDVAEPPIRIFMLLAGLPEDIPHLKYL 180  
Db 121 MPVVDKLENRIQELACSLIESLRQGCNFTEDVAEPPIRIFMLLAGLPEDIPHLKYL 180  
QY 181 TDQMPDGSMTFAEAKAALDYLIPIIEQRQKPGTDALSIIVANGQVNGRPITSDEAKR 240  
Db 181 TDQMPDGSMTFAEAKAALDYLIPIIEQRQKPGTDALSIIVANGQVNGRPITSDEAKR 240  
QY 241 MFCLLVGLDVTVNF-SFSMEFLAKSPEHRQELIERPERIPACBELLRRFSLVADGRI 300  
Db 241 MFCLLVGLDVTVNF-SFSMEFLAKSPEHRQELIERPERIPACBELLRRFSLVADGRI 300

QY 301 LTSDEYEHGVLKKGQDILLPQMLSGLDERKNACPMHVDTSRQKVSHTTTFGHSHLCLGQ 360  
Db 301 LTSDEYEHGVLKKGQDILLPQMLSGLDERKNACPMHVDTSRQKVSHTTTFGHSHLCLGQ 360  
QY 361 HLAARREIIVTLKEMWLRIPDPSIAPGAQIOHKSGIVSGVQALPLVMDPATTKAV 414  
Db 361 HLAARREIIVTLKEMWLRIPDPSIAPGAQIOHKSGIVSGVQALPLVMDPATTKAV 414  
RESULT 5  
US-10-314-657-37  
; Sequence 37, Application US/10314657  
; Publication No. US20030175889A1  
; GENERAL INFORMATION:  
; APPLICANT: SHEN, Ben  
; APPLICANT: CHENG, Yi-Qiang  
; APPLICANT: TANG, Gong-Li  
; TITLE OF INVENTION: Discrete Acyltransferases Associated with Type I Polyketide  
; TITLE OF INVENTION: Syntheses and Methods of Use  
; FILE REFERENCE: 054030-0021  
; CURRENT APPLICATION NUMBER: US/10/314,657  
; CURRENT FILING DATE: 2002-12-03  
; PRIOR APPLICATION NUMBER: PCT/US02/08937  
; PRIOR FILING DATE: 2002-03-22  
; PRIOR APPLICATION NUMBER: US 60/278,935  
; PRIOR FILING DATE: 2001-03-26  
; NUMBER OF SEQ ID NOS: 214  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 37  
; LENGTH: 399  
; TYPE: PRT  
; ORGANISM: Streptomyces atroolivaceus  
US-10-314-657-37

Query Match 77.1%; Score 372; DB 14; Length 399;

Best Local Similarity 28.1%; Pred. No. 1.8e-28;  
Matches 110; Conservative 70; Mismatches 176; Indels 36; Gaps 9;  
QY 38 VQEAWAYLQSNVDPDLVWTRCNGGHWIATRGQIREAYEDYRHFSSCPPIPREAGE--- 94  
Db 21 IHKFAELRETDPLARVLPYGGEGWVTR-----YDDVRAANSDFRFSQAQIGEDTP 73  
QY 95 -----AYDFIPTSDMPPEQORFRALANQVVGMPVDKLENRIQELACSLIESLRPQO 147  
Db 74 RTTPLARRSDTI-LSLDPPEHTRRLRLSKAFTARRMGAMQSLFELFAGLIDGVERTGH 132  
QY 148 -CNFTEDYAEPPPIRIFMLLAGLPEDIPHLKYLTDQWTRPDGSMTFEAKR-----A 199  
Db 133 PADIVRLAQPTTAVICRLGVPEYDGRFQHWSEVI-----MSTTAYSKEEAVSADMS 187  
QY 200 LYDYLIPITIEQRQKPGTDALSIIVANGQVNGRPITSDEAKRMFGLLWGLDVTVNFSLF 259  
Db 188 IRAYLADLVSAARAAPHDDLGLVLSARDDDRLTEDELITFTGVTLLVAGHETSALQLGN 247  
QY 260 SWEFLAKSPEHRQELIERPERIPACBELLRRFSL---VADGRILTSDFEHGVLKKG 316  
Db 248 MVYALLTHEDQSLSLREQPELLPRAVEELLRFVPLGNGVGNARIALDELSVGGTVRAGE 307  
QY 317 QILLPQMLSGLDERKNACPMHVDTSRQKVSHTTTFGHSHLCLGQLARREILVTLKEMLT 376  
Db 308 GVAAAAVNANDPRAFDPPDRLDITREKNPHLAGHGAHYCLGACLARVELKVAIGELLE 367  
QY 377 RIPDPSIA-PCAQIOHKSGIVSGVQALPLVW 406  
Db 368 RFGRLAVPADQVWKTGGTFRGPQRLFIAM 399

## RESULT 6

US-10-214-446-50  
; Sequence 50, Application US/10214446  
; Publication No. US20030180742A1  
; GENERAL INFORMATION:  
; APPLICANT: Weiner, David

```

; APPLICANT: Burk, Mark J.
; APPLICANT: Hitchman, Tim
; APPLICANT: Pujol, Catherine
; APPLICANT: Richardson, Toby
; APPLICANT: Short, Jay M.
; TITLE OF INVENTION: P450 ENZYMES, NUCLEIC ACIDS ENCODING
; TITLE OF INVENTION: THEM AND METHODS OF MAKING AND USING THEM
; FILE REFERENCE: 09010-500001
; CURRENT APPLICATION NUMBER: US/10/214,446
; CURRENT FILING DATE: 2002-08-05
; PRIOR APPLICATION NUMBER: US 60/309,497
; PRIOR FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 50
; LENGTH: 404
; TYPE: PRT
; ORGANISM: Bacterial
US-10-214-446-40

Query Match      16.8%; Score 365.5; DB 14; Length 404;
Best Local Similarity 29.8%; Pred. No. 8.1e-28;
Matches 96; Conservative 64; Mismatches 158; Indels 33; Gaps 5;

Qy 60 GGWIAIRGQIRRAYEDYRHFSCEPPIPREAGEAYDFIPTM-----103
Db 32 GRWFLPHHADIRIALDERFSAS-----RKAGGFVNFPAEVRPEFARNALSRWIV 85
Qy 104 --DPPEQRORFALANQVGMVVDKLENRIQELACSLIESLRPQCCNFTEDYAEPPPIR 161
Db 86 LHDQPEHRQLRQMQCGFTRLITMTPEKIQRCVDELDIAFVXRGSTEFMTYAHPPFAK 145
Qy 162 IFMLLAGLPEDIEHLXYLTDMTRPGSM-----TFRAAKE---ALVDYLIPILIEQRQ 213
Db 146 VIAEMLVNPDYFAFVVMSEDLNFAQSLRPTLEMFRAAQGLLAMDYFARLLPERE 205
Qy 214 KPGTDA-SIVANGQVNGRPITSDKAEMFGLLVGLLTVVNFLSFSEFLAKSPHEHQE 273
Db 206 NPGDVLVSLLSAEGEEMTAQVLANCTQIIIVAGHETTRNLVANGVELLRYPEQAL 265
Qy 274 LIERPERIPACELLRRFS-LVADGRILTSYDEPHGVOLKKGQDILLPOMLSGLDERKN 332
Db 266 LESRPELMPSAVREIMRPESPLOFIRRVAREDFEFGGAEVREGDGLVILMGSANRDPFA 325
Qy 333 ACMEHVDESRQKVGSHTTGHGSHLCLGQHLARE-IIVTLKEWLTTRIPDFSI 383
Db 326 DDDTDFLTRNPTGHLAFGWGPHVCVGAALAELEGQVSEFRILLDRPLGLEL 376

RESULT 7
US-10-214-446-40
; Sequence 40, Application US/10214446
; Publication No. US20030180742A1
; GENERAL INFORMATION:
; APPLICANT: Weiner, David
; APPLICANT: Burk, Mark J.
; APPLICANT: Hitchman, Tim
; APPLICANT: Pujol, Catherine
; APPLICANT: Richardson, Toby
; APPLICANT: Short, Jay M.
; TITLE OF INVENTION: P450 ENZYMES, NUCLEIC ACIDS ENCODING
; TITLE OF INVENTION: THEM AND METHODS OF MAKING AND USING THEM
; FILE REFERENCE: 09010-500001
; CURRENT APPLICATION NUMBER: US/10/214,446
; CURRENT FILING DATE: 2002-08-05
; PRIOR APPLICATION NUMBER: US 60/309,497
; PRIOR FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 40
; LENGTH: 404
; TYPE: PRT
; ORGANISM: Bacterial

US-10-214-446-40
Query Match      16.8%; Score 365.5; DB 14; Length 404;
Best Local Similarity 29.8%; Pred. No. 8.1e-28;
Matches 118; Conservative 58; Mismatches 189; Indels 31; Gaps 10;

Qy 31 PSNLSAGVQBAWVQLQSNVPLDLVWTRCNGGH--WIAIRGQIRRAYEDYRHFSCEPPI 88
Db 20 PFNEADGISLADAYEAREEQGLLRVMAYGEPAWLATR-----YADARLVLDGRFES 72
Qy 89 -----PREAGEAYDFIPTSDPPPEQORFALANQVGMVVDKLENRIQELACSLI 139
Db 73 RAEGARHDEPRQSEGRDGLSMDPPDHTRLRTLVAKAFTMHQVEKLRPAVRELLADELI 132
Qy 140 ESLRPOG-QCNFTEDYAEPPPIRIFMLLAGLPEDIEHLXYLTDMTRPGDSMT---PAE 195
Db 133 DQWATGAPVDLVEEPALPVPVGVICQLLVPEDRPRFRASD-AALSTSLTAEEEDA 191
Qy 196 AKEALYDYLIPILIEQRQKPGTDAISIVANGQVNGRPITSDKAEMFGLLVGLLTVV 255
Db 192 NQEEELRAYMGLIEDHRARPREDLITGLIEARDRDRDLTEQELVDLCVGLVAGHETTAT 251
Qy 256 FLFSMEFLAKSPHEHQELIERPERIPACELLRFSIVADG----RLTSDYSEHGVO 311
Db 252 QIPNFVVTLLDRPQMNRLRECPSELVPTAVBELM-RFVPLSGSGAFPRVATEDVVGCEL 310
Qy 312 LKKGQDILLPOMLSGLDERKNACPMHVDFSRQKVSHTTFFGESSHLCGQHLARREIIVTL 371
Db 311 VRAGEPVLVGAANRDPARFDAQELDLAREGQHLGFGHGVHCLGAPLARLELQAL 370
Qy 372 KEWLTTRIPDFSIAGAQIOHKSQ-IVSGVQALPIVW 406
Db 371 GALLRLRLGLRIA--GDIEWKTQMLVRGPRTLPGVM 404

RESULT 8
US-10-156-761-14997
; Sequence 14997, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 14997
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-14997

Query Match      16.8%; Score 365; DB 14; Length 416;
Best Local Similarity 29.1%; Pred. No. 9.5e-28;
Matches 104; Conservative 56; Mismatches 181; Indels 16; Gaps 8;

Qy 63 WIATRGQIRRAYEDYRHFSCEPPIPR--EAGEAYDFIPT-----SMDPPPEQORFAL 114
Db 63 WYVTGHAARALLSDQRLSSDRT--LPRFPATERFEAVTRRVALGVDPPHETQRM 120
Qy 115 ANQVGMVVDKLENRIQELACSLIESLRPQG-QCNFTEDYAEPPPIRIFMLLAGLPED 173
Db 121 LVFSFTLKRAALRPRIQETVDGLLDAAEQGPALVSAFALPLPSWICALGYPYAD 180
```

```

QY 174 IPHLYTQXTRPDGSMWFAEAKYALYLIPIIEQRQKPGTDALSIIVANGQNGRPI 233
Db 181 HDPFESQRRLRGFGIAEAGDARADDDYLYALIDRKKEPGDGLLDLIQEQLNRGT 240
QY 234 TSDAKYFGLLVGLDVTNFIISFMSFLAKSPEHRQELIERPERIPAAACEILLRFS 293
Db 241 DRALVSAJLLLAGHETTANISLSGFTTLRPEQLAEIRASPGMLPAVAEILL-RFL 299
QY 294 JVADG--RIITSDYEFHGVQLKGDQILLPQMLSGDERKNACPMHVDIFSQRKSHHTFG 351
Db 300 SIADGLLRVATEDIEVAGTT--RADEGVVFATSVINRDAAGFAEPDADLWHSARHVAFG 359
QY 352 HSHLCLGQHLARERIIIVTLKEMTRIPDFSI-APGAQIQKSG-IVSGVQALPLVW 406
Db 360 FGHQCLGQHLARAEIATLGLFRLPGLRLAAPADEIPFKGFTIQGMLELPLVTW 416

```

## RESULT 9

```

US-09-861-289-39
; Sequence 39, Application US/09861289
; Patent No. US2002010897A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/861,289
; CURRENT FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 39
; LENGTH: 416
; TYPE: PRF
; ORGANISM: Streptomyces venezuelae
US-09-861-289-39

```

```

Query Match 16.7%; Score 363.5; DB 9; Length 416;
Best Local Similarity 28.3%; Pred. No. 1.3e-27;
Matches 97; Conservative 62; Mismatches 163; Indels 21; Gaps 7;

QY 76 EDYRHSSECPPIPREAGEAYDIPTSDPPQROFRALANOVGMVGVKLENRIQELA 135
Db 72 KDW--NSTPLTEAEALNHNLES--DPPRHLRLKLVAREFTMRRVVELLRPRVQEV 127

QY 136 CSLIESL--RPOGQCNFTEDYAEPPPIRIFMLLAGIPEDIPHLKYLTDQMTRPDGSMTF 193
Db 128 DGLVDAMLAAPDGRADLMESLAWPLPIVISELLGVPEPDRAAFRWTDADFVDDPAQA 187

QY 194 AEAKYALYLIPIIEQRQKPGTDALSI-IVANGQNGRPIITSDAKRMEGLLVGLDGT 252
Db 188 QTAAEMSGYLSRLIDSKRGQDGLSALVTSDEGSRLSSEELGMAHILLVAGHET 247

QY 253 VVNF--SFMSEFLAKSPEHRQELIERPERIPAAACEILLRFSIVADGRILTSYEF---- 307
Db 248 TVNLJANGMYALLSHPDQALALRADMTLLDGAVEEMLR-----YEGPVESATYRFPVEPV 302

QY 308 --HGVQLKKGDIILPQMLSGDERKNACPMHVDIFSQRKSHHTFGHSHLCLGQHLARR 365
Db 303 DLDGTVIPAGDVLVVLADAHRTPRFPDPHFRDTRDTAGHLAFGHGIFHCIGAPLARL 362

QY 366 ELIIVLKEMLTIPDFS--IAPGAQIQKSGIVSGVQALPLVW 406
Db 363 EARIARALLERCPLDALDVSPGELVWYFNPIMRGLKALPIRW 405

```

## RESULT 10

```

US-09-860-846-39
; Sequence 39, Application US/09860846

```

```

; Patent No. US20020164742A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/860,846
; CURRENT FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 39
; LENGTH: 416
; TYPE: PRF
; ORGANISM: Streptomyces venezuelae
US-09-860-846-39

```

```

Query Match 16.7%; Score 363.5; DB 9; Length 416;
Best Local Similarity 28.3%; Pred. No. 1.3e-27;
Matches 97; Conservative 62; Mismatches 163; Indels 21; Gaps 7;

QY 76 EDYRHSSECPPIPREAGEAYDIPTSDPPQROFRALANOVGMVGVKLENRIQELA 135
Db 72 KDW--NSTPLTEAEALNHNLES--DPPRHLRLKLVAREFTMRRVVELLRPRVQEV 127

QY 136 CSLIESL--RPOGQCNFTEDYAEPPPIRIFMLLAGIPEDIPHLKYLTDQMTRPDGSMTF 193
Db 128 DGLVDAMLAAPDGRADLMESLAWPLPIVISELLGVPEPDRAAFRWTDADFVDDPAQA 187

QY 194 AEAKYALYLIPIIEQRQKPGTDALSI-IVANGQNGRPIITSDAKRMEGLLVGLDGT 252
Db 188 QTAAEMSGYLSRLIDSKRGQDGLSALVTSDEGSRLSSEELGMAHILLVAGHET 247

QY 253 VVNF--SFMSEFLAKSPEHRQELIERPERIPAAACEILLRFSIVADGRILTSYEF---- 307
Db 248 TVNLJANGMYALLSHPDQALALRADMTLLDGAVEEMLR-----YEGPVESATYRFPVEPV 302

QY 308 --HGVQLKKGDIILPQMLSGDERKNACPMHVDIFSQRKSHHTFGHSHLCLGQHLARR 365
Db 303 DLDGTVIPAGDVLVVLADAHRTPRFPDPHFRDTRDTAGHLAFGHGIFHCIGAPLARL 362

QY 366 ELIIVLKEMLTIPDFS--IAPGAQIQKSGIVSGVQALPLVW 406
Db 363 EARIARALLERCPLDALDVSPGELVWYFNPIMRGLKALPIRW 405

```

## RESULT 11

```

US-09-988-384B-39
; Sequence 39, Application US/0988384B
; Publication No. US2003007382A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.536US1
; CURRENT APPLICATION NUMBER: US/09/988,384B
; CURRENT FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: PCT/US99/14398
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 53
; SEQ ID NO 39
; LENGTH: 416
; TYPE: PRF
; ORGANISM: Streptomyces venezuelae
US-09-988-384B-39

```



[illegible][illegible]

Search completed: April 6, 2004, 19:14:31  
Job time : 35.974 secs



```
QY 253 VVNFSLFSMEFLAKSPHQRQLIERPERIPAAACELLRRFSLVADGRILTSYEF----- 307
Db 248 TVNLIANGMYALLSHDPQALRADMTLLDGAVEMLR-----YEGPVESATYRFPPEVP 302
QY 308 --HGVLKKGQDQILLPQMLSGDLDERKNACPMHVDFSRQKVSHTTFHGSHLCLGQHARR 365
Db 303 DLGDTVIPAGDTVLVLAADAHRTPEFPCERHFDIRDTAGHLAFGHGHCIGAPLARR 362
QY 366 EIIIVTLKEWLTRIPDFS--IAPGAQIOHKSIGVSGVQALPLVW 406
Db 363 EARIARALLERCPLDALDVSPGLVWYVNPMPMIRGLKALPIRW 405

RESULT 2
US-09-105-537-39
; Sequence 39, Application US/09105537A
; Patent No. 6265202
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xie, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438U51
; CURRENT APPLICATION NUMBER: US/09/105.537A
; CURRENT FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version: 3.0
; SEQ ID NO 39
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-105-537-39
```

```
Query Match 16.7%; Score 363.5; DB 3; Length 416;
Best Local Similarity 28.3%; Pred. No. 5.5e-29;
Matches 97; Conservative 62; Mismatches 163; Indels 21; Gaps 7;

QY 76 EDYRHFSECFPIPREAGEAYDFIPTSDPPEQRFALANQVGVMPVVKLENRIQELA 135
Db 72 KDMR--NSTTPTTEAAALNNMLES--DPPRHTLRKLVAEFTMRVRLPRVQEV 127
QY 136 CSLIESI--RPQGCNFTEDYAEPPPIRIFMLLAGLPEDIPHLKYLTDQMTDPGSMTF 193
Db 128 DGLVDAMLAAPDGRADIMESLAWPLPITVISSELLGVPEPDRAAEFRVWTDVFPDDPAQA 187
QY 194 AEAEKALVDYLPIIEQRQKPGTDAIS--IVANGQVNGRPITSDRAKMFGLLLVGGLD 252
Db 188 QTAMAEMSGYLSRLDSKRGQDGLLSALVRTSDEGSRUTSELLGMAHILLVAGHET 247
QY 253 VVNFSLFSMEFLAKSPHQRQLIERPERIPAAACELLRRFSLVADGRILTSYEF----- 307
Db 248 TVNLIANGMYALLSHDPQALRADMTLLDGAVEMLR-----YEGPVESATYRFPPEVP 302
QY 308 --HGVLKKGQDQILLPQMLSGDLDERKNACPMHVDFSRQKVSHTTFHGSHLCLGQHARR 365
Db 303 DLGDTVIPAGDTVLVLAADAHRTPEFPCERHFDIRDTAGHLAFGHGHCIGAPLARR 362
QY 366 EIIIVTLKEWLTRIPDFS--IAPGAQIOHKSIGVSGVQALPLVW 406
Db 363 EARIARALLERCPLDALDVSPGLVWYVNPMPMIRGLKALPIRW 405
```

```
RESULT 3
US-09-141-908-13
; Sequence 13, Application US/09141908
; Patent No. 6503741
; GENERAL INFORMATION:
; APPLICANT: Ashley, Gary
; APPLICANT: Betlach, Melanie C.
; APPLICANT: Betlach, Mary
; APPLICANT: McDaniel, Robert
; APPLICANT: Tang, Li
```

```
; TITLE OF INVENTION: Combinatorial Polyketide Libraries Produced Using a
; FILE OF INVENTION: Modular PKS Gene Cluster as Scaffold
; CURRENT APPLICATION NUMBER: US/09/141.908
; CURRENT FILING DATE: 1998-08-28
; EARLIER APPLICATION NUMBER: CIP OF 09/073.538
; EARLIER FILING DATE: 1998-05-06
; EARLIER APPLICATION NUMBER: CIP OF 08/846.247
; EARLIER FILING DATE: 1997-04-30
; EARLIER APPLICATION NUMBER: PROV. 60/376.919
; EARLIER FILING DATE: 1998-03-05
; EARLIER APPLICATION NUMBER: PROV. 60/387.080
; EARLIER FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-141-908-13

Query Match 16.7%; Score 363.5; DB 4; Length 416;
Best Local Similarity 28.3%; Pred. No. 5.5e-29;
Matches 97; Conservative 62; Mismatches 163; Indels 21; Gaps 7;

QY 76 EDYRHFSECFPIPREAGEAYDFIPTSDPPEQRFALANQVGVMPVVKLENRIQELA 135
Db 72 KDMR--NSTTPTTEAAALNNMLES--DPPRHTLRKLVAEFTMRVRLPRVQEV 127
QY 136 CSLIESI--RPQGCNFTEDYAEPPPIRIFMLLAGLPEDIPHLKYLTDQMTDPGSMTF 193
Db 128 DGLVDAMLAAPDGRADIMESLAWPLPITVISSELLGVPEPDRAAEFRVWTDVFPDDPAQA 187
QY 194 AEAEKALVDYLPIIEQRQKPGTDAIS--IVANGQVNGRPITSDRAKMFGLLLVGGLD 252
Db 188 QTAMAEMSGYLSRLDSKRGQDGLLSALVRTSDEGSRUTSELLGMAHILLVAGHET 247
QY 253 VVNFSLFSMEFLAKSPHQRQLIERPERIPAAACELLRRFSLVADGRILTSYEF----- 307
Db 248 TVNLIANGMYALLSHDPQALRADMTLLDGAVEMLR-----YEGPVESATYRFPPEVP 302
QY 308 --HGVLKKGQDQILLPQMLSGDLDERKNACPMHVDFSRQKVSHTTFHGSHLCLGQHARR 365
Db 303 DLGDTVIPAGDTVLVLAADAHRTPEFPCERHFDIRDTAGHLAFGHGHCIGAPLARR 362
QY 366 EIIIVTLKEWLTRIPDFS--IAPGAQIOHKSIGVSGVQALPLVW 406
Db 363 EARIARALLERCPLDALDVSPGLVWYVNPMPMIRGLKALPIRW 405

RESULT 4
US-09-657-440-18
; Sequence 18, Application US/09657440
; Patent No. 6509455
; GENERAL INFORMATION:
; APPLICANT: Ashley, Gary
; APPLICANT: Betlach, Melanie C.
; APPLICANT: Betlach, Mary C.
; APPLICANT: McDaniel, Robert
; APPLICANT: Tang, Li
; TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 300622002120
; CURRENT APPLICATION NUMBER: US/09/657.440
; CURRENT FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: 09/320.878
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: CIP OF 09/141.908
; PRIOR FILING DATE: 1998-08-28
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 416
; TYPE: PRT
```





TEPPERMAN, JAMES M.  
; TITLE OF INVENTION: EXPRESSION OF HERBICIDE METABOLIZING  
; CYTOCHROMES  
; NUMBER OF SEQUENCES: 19  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/569,781  
; FILING DATE: 23-AUG-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 464,499  
; FILING DATE: 12-JAN-1990  
; APPLICATION NUMBER: 405,605  
; FILING DATE: 11-SEP-1989  
; SEQ ID NO.6:  
; LENGTH: 406  
5212296-6

Query Match 15.1%; Score 328; DB 6; Length 406;  
Best Local Similarity 26.3%; Pred. No. 2.5e-25;  
Matches 104; Conservative 74; Mismatches 193; Indels 24; Gaps 11;  
Matches 104; Conservative 74; Mismatches 193; Indels 24; Gaps 11;

QY 31 PSNLSAGVQ--EAMVQLQESNVDLVWTRCNGGH-WIATRGQLIREAYEDYR----- 79  
Db 17 ESNRSCPYQLPDGYAQLRDTPGPHRVTLVDGRQAWVTTKHEARKLLGSDPRLSSNRITD 76  
QY 80 HESSECPFIP--REAGEAYDFIPTSDPPEQORFRALANQVGVVVDKLENRIQELACS 137  
Db 77 NFATSPREFAVRESPOAF----IGLDPPEHGTERRMTISEFTVKRIKGRPEVEVVG 132  
QY 138 LIESLRPQG--QCNFTEYAEPPFIRIPMLLAGUPEEDIPHLKYLTDQMTDPGSMTFEAE 196  
Db 133 FLDEMLAGPTADLVSQFALPVSMTVCRLLGVYADHEFFQASKRLVQSTDAQSALTA 192  
QY 197 KEALYDYLIPILQRCQKRGTDAL-STVANGQNGRPITSDCAKMFGLLVGLDVTVN 255  
Db 193 RNLGAGYLDGLITQFQEPGAGLVGALVADQLANGE-IDREELSTAMLLLIAGHETTAS 251  
QY 256 FLGFSMEFLAKSPEHRQELTERPERIPAAEELLRRFSL--VADGRILTSDFEFGVQLK 313  
Db 252 MTSLSVITLLDHEQYAAALRADSLVPSGAEELLRLYLAIADIAGRGVATADIEVEGLIR 311  
QY 314 KGPQILPQWGLDERKACPMHVDPSRKVSHHTTGHSHLCLGQHLARREIIVTKE 373  
Db 312 AGEGVIVNNSIANRGDVTVEDPOALDIHRSARHHLAFGFGVHQCGLQGLQEARLELEVLNA 371  
QY 374 WLTRIPDFTA--GCAQTHQKSG-IVSGVQALPLVW 406  
Db 372 LMDRVEPTLRVAVPEQVLVLRPGTTIQGVNELPVTM 406

RESULT 8  
US-08-765-907A-10  
; Sequence 10, Application US/08765907A  
; Patent No. 6352839  
; GENERAL INFORMATION:  
; APPLICANT: BLANC, Veronique  
; APPLICANT: THIBAUT, Denis  
; APPLICANT: BAWAS-JACQUES, Nathalie  
; APPLICANT: BLANCHE, Francis  
; APPLICANT: COUZET, Joel  
; APPLICANT: BARRIERE, Jean-Claude  
; APPLICANT: DEBUSSCHE, Laurent  
; APPLICANT: FAMECHON, Alain  
; APPLICANT: PARIS, Jean-Marc  
; APPLICANT: DUTRUC-ROSSET, Gilles  
; TITLE OF INVENTION: Streptogramins And Method For Preparing Same By  
; TITLE OF INVENTION: Mutasynthesis  
; FILE REFERENCE: Streptogramin genes  
; CURRENT APPLICATION NUMBER: US/08/765,907A  
; CURRENT FILING DATE: 1997-03-20  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentID ver. 2.0  
; SEQ ID NO 10  
; LENGTH: 399

TYPE: PRT  
; ORGANISM: Streptomyces pristinaespiralis  
US-08-765-907A-10

Query Match 14.9%; Score 324.5; DB 4; Length 399;  
Best Local Similarity 29.0%; Pred. No. 5.7e-25;  
Matches 106; Conservative 59; Mismatches 159; Indels 41; Gaps 12;

QY 74 AYEDYRH-----FSSSECPFIPREAGEAYDFIPTSDPPEQORFRALANQVGV 121  
Db 36 AFHVERHADVLTVASDPGVYSSQLRSPGSQALSEQLSVIDPFMHTLRRLVSQAFTP 95  
QY 122 PVVDKLENRIQELACSILSRPQGC--NETEDYAEPPFIRIPMLLAGUPEEDIPHLKYL 180  
Db 96 RTVADLEPRVTELQGLLDV--DGDFTDLVADFAYPLVIVIAELLGVPPADRTLFRSW 153  
QY 181 TDQWTR-----PDGSMTFEAKKALYDYLIPILQRCQKRGTDALSIWA 224  
Db 154 SDRMLQWQVADPADMQFGDDADEYQRLVKEPMKAMHAYLHDVTDREARFANDLISALV 213  
QY 225 NGOVNGRPITSDCAKMFGLLVGLDVTVNFLSFSMEFLAKSPEHRQELTERPER--IP 282  
Db 214 AARVEGRLDEQVEFGALLMAGHVTSTMLLNTVLCLXDP--RAEAAARADRSCLIP 271  
QY 283 AACBEELR-RFSLVADGRILTSDFEFGVQ--KKGQDQILLPQWLS--GLDERKACPMHVD 340  
Db 272 ALIEEVLRLRPITVMAV--TKDTVLAGTTIPAG--RMVTPSLLSANHDEQVFTDPDHL 330  
QY 341 SRQKVSHTTGHSHLCLGQHLAREIIVLKEWLTRIPDPSIAPGAQIQ--HKSGIVSGV 399  
Db 331 AREG--RQIAFGHGHYCLGAPLARLEGRIALALFDRPPDFSP--DGAKIRVHRDGLF--GV 388  
QY 400 QALPL 404  
Db 389 KNLPL 393

RESULT 9  
US-09-385-028-12  
; Sequence 12, Application US/09385028  
; Patent No. 6232106  
; GENERAL INFORMATION:  
; APPLICANT: Susan E. Jensen  
; APPLICANT: Kwamena A Aidoo  
; APPLICANT: Ashish S. Paraskar  
; TITLE OF INVENTION: DNA Sequence Encoding Enzymes of Clavulanic  
; Patent No. 6232106  
; TITLE OF INVENTION: Acid Biosynthesis  
; NUMBER OF SEQUENCES: 25  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: JACOBSON, PRICE, HOLMAN & STERN, PLLC  
; STREET: The Jenifer Building, 400 Seventh Street, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.3, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/385,028  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/790,462  
; FILING DATE: 29-JAN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: D. Douglas Price  
; REGISTRATION NUMBER: 24,514  
; REFERENCE/DOCKET NUMBER: 14-8/P57452US2  
; TELECOMMUNICATION INFORMATION:

```
; TELEPHONE: (202 638-6666
; TELEFAX: (202) 39305350
; TELEX: RCA 248593 IDEA UR
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 409 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-385-028-12

Query Match          14.8%; Score 322; DB 3; Length 409;
Best Local Similarity 28.7%; Pred. No. 1.1e-24;
Matches 121; Conservative 60; Mismatches 186; Indels 54; Gaps 18;

QY 16 PHVEHLVDFDMMYNPSNLGAGVOEAWAVLOESNVPDLVWTRCNCGGH-WIATRGQLIREA 74
DB 14 PAYEMGRVCPVD---PPQLAGLSRQKAASRVT-----LW---DGSQVWLVTSHAGARAV 62
QY 75 YEDYRHFS-SECPFIP-----REAGEAYDFIPTSMDDPPEQCFRA-----LANOV 118
DB 63 LGDRRTAVTSAPGFPMLTRISQLVNRPESASFI--RMDDPQHSRLRSLMTRDFLARRA 120
QY 119 VGM-PVVDKLENRIQELACSLIESLRPGQCNFTEDYAEPPPIRIFMLLAGLPEDIPHL 177
DB 121 EALRPVAVREL---LDEILGGLVKGPER--VDLVAGLTTPVPSRVITLLFGAGDDRREFI 174
QY 178 K-----YLTQMTRPDGSMTFAEAKALDYLIPIIEQRQKPGTDAISIVANGQVNGRPI 233
DB 175 EDRSAVLIDRGYTP---QVAKARDELGDYLRLEIVEERIENPGTDLISRLVIDQVRPGHL 231
QY 234 TSEAKEMFGLLVGGLDVTVMF--SFSMEFLAKSPEHQELIEREPIPAACEELLRRFS 293
DB 232 RVEEMVPMCRLLLVAGHGTTTSQASLSLLTDPDELAGRLTEDPALLPKAVEELLRPHS 291
QY 294 LVADG--RLITSYEFHGVCLKKGQDQILLPQMLSGDLERKNACPMHVDFFSRQKVSHTTFF 351
DB 292 IVQNGLARAAVEDVQDDVLIRAGEGVVLSLSAGNRDETVPDPDRVDVDRDARRHLAFG 351
QY 352 HGHSLCLGQHLAR---REIIVTLKEWLTRIPDSIA-PGAQIOHKSGIVS-GVQALPLVM 406
DB 352 HGMHQCLGQWLARVELEILLAAVLRW---PGARLAVPFEELDFRHEVSSVGLGALPVTM 408
QY 407 D 407
DB 409 Z 409

RESULT 10
US-09-726-614-12
; Sequence 12, Application US/09726614
; Patent No. 6514735
; GENERAL INFORMATION:
; APPLICANT: Susan E. Jensen
; APPLICANT: Kwamena A. Aidoo
; APPLICANT: Ashish S. Paradkar
; TITLE OF INVENTION: DNA Sequence Encoding Enzymes of Clavulanic
; Patent No. 6514735
; TITLE OF INVENTION: Acid Biosynthesis
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: JACOBSON, PRICE, HOLMAN & STERN, PLLC
; STREET: The Jennifer Building, 400 Seventh Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0.0, Version #1.30 (EPO)
```

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/726,614
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/790,462
; FILING DATE: 29-JAN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: D. Douglas Price
; REGISTRATION NUMBER: 24,514
; REFERENCE/DOCKET NUMBER: 1418/P57452US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 39305350
; TELEX: RCA 248593 IDEA UR
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 409 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-726-614-12

Query Match          14.8%; Score 322; DB 4; Length 409;
Best Local Similarity 28.7%; Pred. No. 1.1e-24;
Matches 121; Conservative 60; Mismatches 186; Indels 54; Gaps 18;

QY 16 PHVEHLVDFDMMYNPSNLGAGVOEAWAVLOESNVPDLVWTRCNCGGH-WIATRGQLIREA 74
DB 14 PAYEMGRVCPVD---PPQLAGLSRQKAASRVT-----LW---DGSQVWLVTSHAGARAV 62
QY 75 YEDYRHFS-SECPFIP-----REAGEAYDFIPTSMDDPPEQCFRA-----LANOV 118
DB 63 LGDRRTAVTSAPGFPMLTRISQLVNRPESASFI--RMDDPQHSRLRSLMTRDFLARRA 120
QY 119 VGM-PVVDKLENRIQELACSLIESLRPGQCNFTEDYAEPPPIRIFMLLAGLPEDIPHL 177
DB 121 EALRPVAVREL---LDEILGGLVKGPER--VDLVAGLTTPVPSRVITLLFGAGDDRREFI 174
QY 178 K-----YLTQMTRPDGSMTFAEAKALDYLIPIIEQRQKPGTDAISIVANGQVNGRPI 233
DB 175 EDRSAVLIDRGYTP---QVAKARDELGDYLRLEIVEERIENPGTDLISRLVIDQVRPGHL 231
QY 234 TSEAKEMFGLLVGGLDVTVMF--SFSMEFLAKSPEHQELIEREPIPAACEELLRRFS 293
DB 232 RVEEMVPMCRLLLVAGHGTTTSQASLSLLTDPDELAGRLTEDPALLPKAVEELLRPHS 291
QY 294 LVADG--RLITSYEFHGVCLKKGQDQILLPQMLSGDLERKNACPMHVDFFSRQKVSHTTFF 351
DB 292 IVQNGLARAAVEDVQDDVLIRAGEGVVLSLSAGNRDETVPDPDRVDVDRDARRHLAFG 351
QY 352 HGHSLCLGQHLAR---REIIVTLKEWLTRIPDSIA-PGAQIOHKSGIVS-GVQALPLVM 406
DB 352 HGMHQCLGQWLARVELEILLAAVLRW---PGARLAVPFEELDFRHEVSSVGLGALPVTM 408
QY 407 D 407
DB 409 Z 409

RESULT 11
US-09-385-040-12
; Sequence 12, Application US/09385040
; Patent No. 6589775
; GENERAL INFORMATION:
; APPLICANT: Susan E.
; APPLICANT: Aidoo, Kwamena A.
; APPLICANT: Paradkar, Ashish S.
; TITLE OF INVENTION: DNA SEQUENCE ENCODING ENZYMES OF CLAVULANIC ACID
; TITLE OF INVENTION: BIOSYNTHESIS
; FILE REFERENCE: 09/385,040
; CURRENT APPLICATION NUMBER: US/09/385,040
```

```

; CURRENT FILING DATE: 1999-08-30
; PRIOR APPLICATION NUMBER: US 08/790,452
; PRIOR FILING DATE: 1997-01-29
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 409
; TYPE: PRT
; ORGANISM: Streptomyces clavuligerus
US-09-385-040-12

Query Match      14.8%; Score 322; DB 4; Length 409;
Best Local Similarity 28.7%; Pred. No. 1,1e-24;
Matches 121; Conservative 60; Mismatches 186; Indels 54; Gaps 18;

QY 16 PHVHELVDFDMYFNSLSAGVQEAVALQESNVECLVWTRNGGH-WIATRGQILIREA 74
Db 14 PAYPMHVCVVD---PPQLAGLRSQKAASRVT---LW---DGSQVWLVTSHAGARV 62
QY 75 YEDYRHFS-SECFPEF-----REAGEAYDFI27SMDDPEQRFRA-----LANQV 118
Db 63 LGRDRRTAVTSAPGPFMLTSTLSQVRANPESASFI--RMDDPQHSRLRSLMTRDFLAR 120
QY 119 VGM-PVWVKLENRIOLACSLIESLRPQGCNFTEDYAEPPRIIFMLLAGLPEEDIPHL 177
Db 121 EALRPVAVREL---LDLILGLVKGER---VDLVAGLITVPVSRVITLLFGAGDGRREFI 174
QY 178 K-----YLTQMTRPDGSMTFAEAKALYLYLPIIQRQKPGTDAISIVANGQVNGRPI 233
Db 175 EDRSAVLIDRGYTFE---QVAKARDELGYLRELVBEERIEENPGTDLSIAZVIDQVRPHL 231
QY 234 TSDEAKEMFGLIAGVGLDVTWNFLSFSMEFLAKSPHQRLIERPERIPAAACEILLRFS 293
Db 232 RVEEMVPMKRLLEVAGHGTT--SGASLSLSLLTDPELAGHLETDPAALLPAVVELLRPHS 291
QY 294 LVADG--RLTSDYEFHGVQKKGQDILLPQMLSGLDERKXNACPMRVDFSRQKVSHTTFG 351
Db 292 IVQNGLARAAVEDVQDDVLIRAGEGVLSLSAGNRDETVEFPDVRVDVDRDARRHLAFG 351
QY 352 HSHLCLGQHLAR---REIIVTKEMLTRIPDFSTIA-PGAQ-CHKSGIVS-GVQALPLVW 406
Db 352 HGMHQLGOWLARVELEELAAVLRW---PGARLAVFFELDFRHEVSSYGLGALPVTW 408
QY 407 D 407
Db 409 Z 409

```

```

RESULT 12
US-08-102-863-11
; Sequence 11, Application US/08102863
; Patent No. 5466590
; GENERAL INFORMATION:
; APPLICANT: SARIASLANI, SIMA
; TITLE OF INVENTION: CONSTITUTIVE
; TITLE OF INVENTION: EXPRESSION OF P450SOY
; TITLE OF INVENTION: AND FEREDOXIN-SOY IN
; TITLE OF INVENTION: STREPTOMYCES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: USA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0,
; SOFTWARE: Version #1.25

```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/102,863
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/807,001
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: GALLEGO, R. THOMAS
; REGISTRATION NUMBER: 32,692
; REFERENCE/DOCKET NUMBER: CR-9000
; TELEPHONE: 302-892-7342
; TELEFAX: 302-892-7949
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 412 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-102-863-11

Query Match      14.4%; Score 314; DB 1; Length 412;
Best Local Similarity 27.2%; Pred. No. 7.3e-24;
Matches 84; Conservative 56; Mismatches 163; Indels 6; Gaps 5;

QY 103 MDPEQRFALANQVGMVVDKLENRIOLACSLIESLRPOG-QCNFTEDYAEPPRI 161
Db 105 VDDPEHNTQRMILPTFSVKRIGALREIQTVDRLDAMEROGPPAELVSFAFALPFSM 164
QY 162 IFMLLAGLPEEDIPHLKYLTDQMTRPDGSMTFAEAKALYLYLPIIQRQKPGTDAIS 221
Db 165 VICALLGVYADHAFPEERSQRLLRGEGADVNARDELEBYLGALIDRKAEPGCGLLD 224
QY 222 IVANGQVNGRPTSDENRVMGILLVGLDVTNVNLSFSMEFLAKSPHQRLIERPERI 281
Db 225 ELTHRDHPDGPVDRQELVAFVILLIAGHETTANMISLGTFTLLSHPEQLAALRAGTST 284
QY 282 PAACEILLRFSLVADG--RLTSDYEFHGVQKKGQDILLPQMLSGLDERKXNACPMHVD 339
Db 285 AVVVEELL-RFLSIAEGLQRLATEDMEVDGATIRKGGVWVFTSLINRDADVFFRAETLD 343
QY 340 FSRQKVSHTTFEGSHLCLGQHLARREIIVTKELWLRIPDFSTIA-PCAQIQHKSIG-IVS 397
Db 344 WDRPARHHLAFGFGVHQCLQNLARALDIAMRTLPERLPGLR-LAVPAHEIRHKPGDTIQ 403
QY 398 GVQALPLVW 406
Db 404 GLLDLPVAV 412

```

```

RESULT 13
PCT-US92-10885-11
; Sequence 11, Application PC/TUS9210885
; GENERAL INFORMATION:
; APPLICANT: SARIASLANI, SIMA
; TITLE OF INVENTION: CONSTITUTIVE
; TITLE OF INVENTION: EXPRESSION OF P450SOY
; TITLE OF INVENTION: AND FEREDOXIN-SOY IN
; TITLE OF INVENTION: STREPTOMYCES
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: USA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch,
; MEDIUM TYPE: 1.0 MB

```



us-09-246-451a-13.rai

Wed Apr 7 09:15:31 2004

```

QY 6 QSNANLAPLPEHVPEHLVEFDXNPSNLSAGVQE-----AWAVLQSNVDPDLVWTRCNG 61
Db 3 QEQANSEIKP-----AEDFKFPAP-----GYAEDPFPAPAESLREA-TFIFYWD--EGR 48
QY 62 HWIATRGQ-----LIREAYDHRFSSECEPFIPIREACEAYDEIPTSMDDPEOR 103
Db 49 SWULTRYHDVSAVFRDERFAVSSEWESSAIP-----ELSDMKYGLFGLPPEDHA 104
QY 110 QFALANQVGMPPVVDKLENRIQELACSLIESLRPOQCNCFTEDYAEPPFIRIFMLLAGL 169
Db 105 RVKXNVNPSFTSRAIDLRLRAEIQTVDQLDARSQOEEDVWRDYASGIPMRAISALKV 164
QY 170 PEEDIPHLKVLTDQMTRPDGSMTFAAKAALYKLP-----I 207
Db 165 PAE-----CDEKERRFGSAT-----ARALGVGLVPOVDEETKLVASVTEGLALLHDV 212
QY 208 IEQRQXP-GTDAISIVANGQVNGRPITSDAKRMFGELLVGLDVTVNFLSFSMEELAK 266
Db 213 LDERRNPLENDVLTMLQAADGSRUSTKELVALNGAITAAGTDTTIIYLIAPAVNLJR 272
QY 267 SPEHQELIERPERIPAAACEELLRRFSLVADG--RILTSDEFFHGVOLKKGQOI--LLPQ 322
Db 273 SPEALELYKABEGLMRNALDEVRFDNILRIGTVFARQDLEYCCASIKKGMVELLIPS 332
QY 323 MLSGLDERKNACPMHVDFSRQKVSHTTFGHSHLCLOHRLARREIIVTLKEWLTIPDPS 382
Db 333 ALR--DGTVFSRPPVDVRRDTGASLAYGRGPHVCPGVSLARLEAEIAGVTIFRRFPENK 390
QY 383 - 383
Db 391 L 391

```

Search completed: April 6, 2004, 18:56:00  
Job time : 15.9588 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OK: protein - protein search, using sw model

Run on: April 6, 2004, 18:53:57 ; Search time 26.1038 Seconds  
(without alignments)  
3108.883 Million cell updates/sec

Title: US-09-246-451A-17

Perfect score: 1508

Sequence: 1 MQLTFFYNVNSCPVNSIVR.....PLTGTGGQIRLNCRVWNSNS 309

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1071772 seqs, 26263353 residues

Total number of hits satisfying chosen parameters: 1071772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

```
1: /cgn2_6/ptodata/1/pubpaa/us07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1608	100.0	309	15 US-10-453-104-17	Sequence 17, Appl
2	1603	99.7	309	14 US-10-274-793-4	Sequence 4, Appl
3	1601	99.6	309	14 US-10-274-793-6	Sequence 6, Appl
4	1586	99.3	309	14 US-10-274-793-18	Sequence 18, Appl
5	1595	99.2	309	14 US-10-274-793-8	Sequence 8, Appl
6	1595	99.2	309	14 US-10-274-793-10	Sequence 10, Appl
7	1594	99.1	309	14 US-10-274-793-12	Sequence 12, Appl
8	1590	98.9	309	14 US-10-274-793-14	Sequence 14, Appl
9	1577	98.1	309	14 US-10-274-793-20	Sequence 20, Appl
10	1577	98.1	309	14 US-10-274-793-24	Sequence 24, Appl
11	1572	97.8	309	14 US-10-274-793-22	Sequence 22, Appl
12	921	57.3	352	12 US-10-424-599-278555	Sequence 278555,
13	912.5	56.7	354	12 US-10-424-599-250229	Sequence 250229,
14	885	55.0	353	12 US-10-424-599-251628	Sequence 251628,
15	874.5	54.4	331	12 US-10-424-599-276196	Sequence 276196,

16	860.5	53.5	333	14 US-10-259-165-18	Sequence 18, Appl
17	860.5	53.5	333	14 US-10-259-165-370	Sequence 370, Appl
18	846.5	52.6	352	12 US-10-424-599-249564	Sequence 249564,
19	836.5	52.0	349	12 US-10-425-114-58420	Sequence 58420, A
20	813	50.6	358	12 US-10-424-599-159302	Sequence 159302,
21	799	49.7	339	12 US-10-425-114-47322	Sequence 47322, A
22	781	48.6	344	14 US-10-289-757-111	Sequence 111, Appl
23	779	48.4	330	12 US-10-424-599-199360	Sequence 199360,
24	770.5	47.9	334	14 US-10-259-165-168	Sequence 168, Appl
25	767.5	47.7	335	16 US-10-389-566-716	Sequence 716, Appl
26	767.5	47.7	335	16 US-10-389-566-2314	Sequence 2314, Ap
27	767.5	47.7	339	14 US-10-259-165-60	Sequence 60, Appl
28	767.5	47.7	339	14 US-10-259-165-396	Sequence 396, Appl
29	767.5	47.7	333	16 US-10-389-566-549	Sequence 549, Appl
30	765.5	47.6	329	16 US-10-389-566-1905	Sequence 1905, Ap
31	758.5	47.2	338	16 US-10-389-566-1496	Sequence 1496, Ap
32	758.5	47.2	338	16 US-10-389-566-1602	Sequence 1602, Ap
33	757.5	47.1	333	16 US-10-389-566-891	Sequence 891, Appl
34	755.5	47.0	334	16 US-10-389-566-2367	Sequence 2367, Ap
35	744.5	46.3	333	16 US-10-389-566-869	Sequence 869, Appl
36	741.5	46.1	316	12 US-10-424-599-279090	Sequence 279090,
37	737.5	45.9	339	16 US-10-389-566-658	Sequence 658, Appl
38	735	45.7	347	12 US-10-424-599-247654	Sequence 247654,
39	733.5	45.6	337	16 US-10-389-566-1239	Sequence 1239, Ap
40	733.5	45.6	337	16 US-10-389-566-2366	Sequence 2366, Ap
41	733	45.6	276	12 US-10-425-114-47525	Sequence 47525, A
42	730.5	45.4	329	16 US-10-389-566-2383	Sequence 2383, Ap
43	728.5	45.3	320	12 US-10-424-599-184544	Sequence 184544,
44	727.5	45.2	315	12 US-10-424-599-246059	Sequence 246059,
45	727.5	45.2	318	14 US-10-174-693-389	Sequence 389, Appl

#### ALIGNMENTS

RESULT 1  
US-10-453-104-17  
; Sequence 17, Application US/10453104  
; Publication No. US20030207345A1  
; GENERAL INFORMATION:  
; APPLICANT: California Institute of Technology;  
; APPLICANT: Frances H. Arnold  
; APPLICANT: Hyun Joo  
; TITLE OF INVENTION: Oxygenase Enzymes and Screening Method  
; FILE REFERENCE: 4358/1B827-US3  
; CURRENT APPLICATION NUMBER: US/10/453,104  
; CURRENT FILING DATE: 2003-06-02  
; PRIOR APPLICATION NUMBER: US 09/661,093  
; PRIOR FILING DATE: 2000-09-13  
; PRIOR APPLICATION NUMBER: US 09/246,451  
; PRIOR FILING DATE: 1999-02-09  
; PRIOR APPLICATION NUMBER: US 60/094,403  
; PRIOR FILING DATE: 1998-07-28  
; PRIOR APPLICATION NUMBER: US 60/106,840  
; PRIOR FILING DATE: 1998-11-03  
; PRIOR APPLICATION NUMBER: US 60/086,206  
; PRIOR FILING DATE: 1998-05-21  
; PRIOR APPLICATION NUMBER: US 60/186,834  
; PRIOR FILING DATE: 1998-11-03  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 17  
; LENGTH: 303  
; TYPE: PRT  
; ORGANISM: Escherichia coli  
US-10-453-104-17

Query Match 100.0%; Score 1608; DB 15; Length 309;  
Best Local Similarity 100.0%; Pred. No. 7.5e-163;  
Matches 309; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQLTFFYNVNSCPVNSIVRDTVNEIUSDPRIAASTRLHFHDCFVNGCDASILNNTT 60  
|||||

Db 1 MQLTPTFYDSCNVSNIVREDTIVNELRSDPRIAASILRLHFHDFCVNGCDASILLNNTT 60  
QY 61 SFTEKDAFGNANSARGFPVIDRMKAAVESACPTVSCADLLTIAAQQSVTLAGGPSWRV 120  
Db 61 SFTEKDAFGNANSARGFPVIDRMKAAVESACPTVSCADLLTIAAQQSVTLAGGPSWRV 120  
QY 121 PLGRDSLOAFDLANANLPAFPFTLPQLKDSFRNVLNRSDDLVALSGGHTFGKNQCRF 180  
Db 121 PLGRDSLOAFDLANANLPAFPFTLPQLKDSFRNVLNRSDDLVALSGGHTFGKNQCRF 180  
QY 181 IMORLYNFSNTGLPDPTLNTTYLQTLRGLCPNLGNLSALVDFDLRTPTIFDNKYVNL 240  
Db 181 IMORLYNFSNTGLPDPTLNTTYLQTLRGLCPNLGNLSALVDFDLRTPTIFDNKYVNL 240  
QY 241 OKGLIQSDQELFSSPDATDIPLVRSFANSTQTFNFVAFVEMDRMGNTPLTGTGGQIRL 300  
Db 241 OKGLIQSDQELFSSPDATDIPLVRSFANSTQTFNFVAFVEMDRMGNTPLTGTGGQIRL 300  
QY 301 NCRVNSNS 309  
Db 301 NCRVNSNS 309  
RESULT 2  
US-10-274-793-4  
; Sequence 4, Application US/10274793  
; Publication No. US20030153042A1  
; GENERAL INFORMATION:  
; APPLICANT: CALIFORNIA INSTITUTE OF TECHNOLOGY;  
; APPLICANT: Frances S. ARNOLD  
; APPLICANT: Zhanglin Lin  
; TITLE OF INVENTION: Expression of Functional Eukaryotic  
; FILE REFERENCE: 3369/1E804-US3  
; CURRENT APPLICATION NUMBER: US/10/274,793  
; PRIOR FILING DATE: 2002-10-21  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: US 09/538,591  
; PRIOR FILING DATE: 2000-03-27  
; PRIOR APPLICATION NUMBER: US 09/247,232  
; PRIOR FILING DATE: 1999-02-09  
; PRIOR APPLICATION NUMBER: US 60/094,403  
; PRIOR FILING DATE: 1998-07-28  
; PRIOR APPLICATION NUMBER: US 60/106,840  
; PRIOR FILING DATE: 1998-11-03  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 4  
; LENGTH: 309  
; TYPE: PRT  
; ORGANISM: Escherichia coli  
US-10-274-793-4

Query Match 99.7%; Score 1603; DB 14; Length 309;  
Best Local Similarity 99.7%; Pred. No. 2.6e-162;  
Matches 309; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MQLTPTFYDSCNVSNIVREDTIVNELRSDPRIAASILRLHFHDFCVNGCDASILLNNTT 60  
Db 1 MQLTPTFYDSCNVSNIVREDTIVNELRSDPRIAASILRLHFHDFCVNGCDASILLNNTT 60  
QY 61 SFTEKDAFGNANSARGFPVIDRMKAAVESACPTVSCADLLTIAAQQSVTLAGGPSWRV 120  
Db 61 SFTEKDAFGNANSARGFPVIDRMKAAVESACPTVSCADLLTIAAQQSVTLAGGPSWRV 120  
QY 121 PLGRDSLOAFDLANANLPAFPFTLPQLKDSFRNVLNRSDDLVALSGGHTFGKNQCRF 180  
Db 121 PLGRDSLOAFDLANANLPAFPFTLPQLKDSFRNVLNRSDDLVALSGGHTFGKNQCRF 180  
QY 181 IMORLYNFSNTGLPDPTLNTTYLQTLRGLCPNLGNLSALVDFDLRTPTIFDNKYVNL 240  
Db 181 IMORLYNFSNTGLPDPTLNTTYLQTLRGLCPNLGNLSALVDFDLRTPTIFDNKYVNL 240

QY 241 OKGLIQSDQELFSSPDATDIPLVRSFANSTQTFNFVAFVEMDRMGNTPLTGTGGQIRL 300  
Db 241 OKGLIQSDQELFSSPDATDIPLVRSFANSTQTFNFVAFVEMDRMGNTPLTGTGGQIRL 300  
QY 301 NCRVNSNS 309  
Db 301 NCRVNSNS 309  
RESULT 3  
US-10-274-793-6  
; Sequence 6, Application US/10274793  
; Publication No. US20030153042A1  
; GENERAL INFORMATION:  
; APPLICANT: CALIFORNIA INSTITUTE OF TECHNOLOGY;  
; APPLICANT: Frances S. ARNOLD  
; APPLICANT: Zhanglin Lin  
; TITLE OF INVENTION: Expression of Functional Eukaryotic  
; FILE REFERENCE: 3369/1E804-US3  
; CURRENT APPLICATION NUMBER: US/10/274,793  
; PRIOR FILING DATE: 2002-10-21  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: US 09/538,591  
; PRIOR FILING DATE: 2000-03-27  
; PRIOR APPLICATION NUMBER: US 09/247,232  
; PRIOR FILING DATE: 1999-02-09  
; PRIOR APPLICATION NUMBER: US 60/094,403  
; PRIOR FILING DATE: 1998-07-28  
; PRIOR APPLICATION NUMBER: US 60/106,840  
; PRIOR FILING DATE: 1998-11-03  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 6  
; LENGTH: 309  
; TYPE: PRT  
; ORGANISM: Escherichia coli  
US-10-274-793-6

Query Match 99.6%; Score 1601; DB 14; Length 309;  
Best Local Similarity 99.4%; Pred. No. 4.2e-162;  
Matches 307; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MQLTPTFYDSCNVSNIVREDTIVNELRSDPRIAASILRLHFHDFCVNGCDASILLNNTT 60  
Db 1 MQLTPTFYDSCNVSNIVREDTIVNELRSDPRIAASILRLHFHDFCVNGCDASILLNNTT 60  
QY 61 SFTEKDAFGNANSARGFPVIDRMKAAVESACPTVSCADLLTIAAQQSVTLAGGPSWRV 120  
Db 61 SFTEKDAFGNANSARGFPVIDRMKAAVESACPTVSCADLLTIAAQQSVTLAGGPSWRV 120  
QY 121 PLGRDSLOAFDLANANLPAFPFTLPQLKDSFRNVLNRSDDLVALSGGHTFGKNQCRF 180  
Db 121 PLGRDSLOAFDLANANLPAFPFTLPQLKDSFRNVLNRSDDLVALSGGHTFGKNQCRF 180  
QY 181 IMORLYNFSNTGLPDPTLNTTYLQTLRGLCPNLGNLSALVDFDLRTPTIFDNKYVNL 240  
Db 181 IMORLYNFSNTGLPDPTLNTTYLQTLRGLCPNLGNLSALVDFDLRTPTIFDNKYVNL 240  
QY 241 OKGLIQSDQELFSSPDATDIPLVRSFANSTQTFNFVAFVEMDRMGNTPLTGTGGQIRL 300  
Db 241 OKGLIQSDQELFSSPDATDIPLVRSFANSTQTFNFVAFVEMDRMGNTPLTGTGGQIRL 300  
QY 301 NCRVNSNS 309  
Db 301 NCRVNSNS 309

RESULT 4  
US-10-274-793-18  
; Sequence 18, Application US/10274793



```

; Publication No. US20030153042A1
; GENERAL INFORMATION:
; APPLICANT: CALIFORNIA INSTITUTE OF TECHNOLOGY;
; APPLICANT: Frances S. ARNOLD
; APPLICANT: Zhonglin LIN
; TITLE OF INVENTION: Expression of Functional Eukaryotic
; TITLE OF INVENTION: proteins
; FILE REFERENCE: 3369/1E804-US3
; CURRENT APPLICATION NUMBER: US/10/274,793
; CURRENT FILING DATE: 2002-10-21
; PRIOR APPLICATION NUMBER: US/09/654,493
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: US 09/538,591
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: US 09/247,232
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: US 60/094,403
; PRIOR FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: US 60/106,840
; PRIOR FILING DATE: 1998-11-03
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Escherichia coli
; US-10-274-793-18

Query Match          99.3%; Score 1596; DB 14; Length 309;
Best Local Similarity 99.4%; Pred. No. 1.4e-161;
Matches 307; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MQLTPTFDNSCPNVSVIRDTIVNELRSDPRIAASILRLHFHDCFVNGCDASILLDNTT 60
Db 1 MQLTPTFDNSCPNVSVIRDTIVNELRSDPRIAASILRLHFHDCFVNGCDASILLDNTT 60
Qy 61 SFTKDAFGNANSARGFPVIDRMKAAVESACPTVSCADLLTAAQOSVTLAGGPSNRV 120
Db 61 SFTKDAFGNANSARGFPVIDRMKAAVESACPTVSCADLLTAAQOSVTLAGGPSNRV 120
Qy 121 PLGRDSLOAFDLANANLPAPFTLPQLKDSFRNVGLNRSSDLVALSGGHTFGKNQCRF 180
Db 121 PLGRDSLOAFDLANANLPAPFTLPQLKDSFRNVGLNRSSDLVALSGGHTFGKNQCRF 180
Qy 181 IMDLNFSNTGLPDPPTINTYLOTLRGLCPNGNLSALVDFDLRTPTIFDKKYVNLEE 240
Db 181 IMDLNFSNTGLPDPPTINTYLOTLRGLCPNGNLSALVDFDLRTPTIFDKKYVNLEE 240
Qy 241 QKGLIQSDQELFSSPDATDTIPLVRSFANSTQTFNAPFVAMDRMGNITPLTGTOGQIRL 300
Db 241 QKGLIQSDQELFSSPDATDTIPLVRSFANSTQTFNAPFVAMDRMGNITPLTGTOGQIRL 300
Qy 301 NCRVNSNS 309
Db 301 NCRVNSNS 309

RESULT 5
US-10-274-793-8
; Sequence 8, Application US/10274793
; Publication No. US20030153042A1
; GENERAL INFORMATION:
; APPLICANT: CALIFORNIA INSTITUTE OF TECHNOLOGY;
; APPLICANT: Frances S. ARNOLD
; APPLICANT: Zhonglin LIN
; TITLE OF INVENTION: Expression of Functional Eukaryotic
; TITLE OF INVENTION: proteins
; FILE REFERENCE: 3369/1E804-US3
; CURRENT APPLICATION NUMBER: US/10/274,793
; CURRENT FILING DATE: 2002-10-21
; PRIOR APPLICATION NUMBER: US/09/654,493
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: US 09/538,591

```

```

; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: US 09/247,232
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: US 60/094,403
; PRIOR FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: US 60/106,840
; PRIOR FILING DATE: 1998-11-03
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Escherichia coli
; US-10-274-793-8

Query Match          99.2%; Score 1595; DB 14; Length 309;
Best Local Similarity 99.0%; Pred. No. 1.8e-161;
Matches 306; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MQLTPTFDNSCPNVSVIRDTIVNELRSDPRIAASILRLHFHDCFVNGCDASILLDNTT 60
Db 1 MQLTPTFDNSCPNVSVIRDTIVNELRSDPRIAASILRLHFHDCFVNGCDASILLDNTT 60
Qy 61 SFTKDAFGNANSARGFPVIDRMKAAVESACPTVSCADLLTAAQOSVTLAGGPSNRV 120
Db 61 SFTKDAFGNANSARGFPVIDRMKAAVESACPTVSCADLLTAAQOSVTLAGGPSNRV 120
Qy 121 PLGRDSLOAFDLANANLPAPFTLPQLKDSFRNVGLNRSSDLVALSGGHTFGKNQCRF 180
Db 121 PLGRDSLOAFDLANANLPAPFTLPQLKDSFRNVGLNRSSDLVALSGGHTFGKNQCRF 180
Qy 181 IMDLNFSNTGLPDPPTINTYLOTLRGLCPNGNLSALVDFDLRTPTIFDKKYVNLEE 240
Db 181 IMDLNFSNTGLPDPPTINTYLOTLRGLCPNGNLSALVDFDLRTPTIFDKKYVNLEE 240
Qy 241 QKGLIQSDQELFSSPDATDTIPLVRSFANSTQTFNAPFVAMDRMGNITPLTGTOGQIRL 300
Db 241 QKGLIQSDQELFSSPDATDTIPLVRSFANSTQTFNAPFVAMDRMGNITPLTGTOGQIRL 300
Qy 301 NCRVNSNS 309
Db 301 NCRVNSNS 309

RESULT 6
US-10-274-793-10
; Sequence 10, Application US/10274793
; Publication No. US20030153042A1
; GENERAL INFORMATION:
; APPLICANT: CALIFORNIA INSTITUTE OF TECHNOLOGY;
; APPLICANT: Frances S. ARNOLD
; APPLICANT: Zhonglin LIN
; TITLE OF INVENTION: Expression of Functional Eukaryotic
; TITLE OF INVENTION: proteins
; FILE REFERENCE: 3369/1E804-US3
; CURRENT APPLICATION NUMBER: US/10/274,793
; CURRENT FILING DATE: 2002-10-21
; PRIOR APPLICATION NUMBER: US/09/654,493
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: US 09/538,591
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: US 09/247,232
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: US 60/094,403
; PRIOR FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: US 60/106,840
; PRIOR FILING DATE: 1998-11-03
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Escherichia coli

```

US-10-274-793-10

Query Match 99.2%; Score 1595; DB 14; Length 309;  
Best Local Similarity 99.0%; Pred. No. 1.8e-161;  
Matches 306; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MQLTPTFYDNCNPNVSNIVRDTIVNELRSDPRIAASILRLHFHDCFVNGCDASILLDNTT 60  
DB 1 MQLTPTFYDNCNPNVSNIVRDTIVNELRSDPRIAASILRLHFHDCFVNGCDASILLDNTT 60  
QY 61 SPRTKDAFGNANSARGFPVIDRMKAAVESACPTVSCADLLTIAAQSVTLAGGFSWRV 120  
DB 61 SPRTKDAFGNANSARGFPVIDRMKAAVESACPTVSCADLLTIAAQSVTLAGGFSWRV 120  
QY 121 PLGRDLSLOAFDLANANLPAPFPTLPOLKDSFRNVLNRSDDLVALSGHTFGKNQCRF 180  
DB 121 PLGRDLSLOAFDLANANLPAPFPTLPOLKDSFRNVLNRSDDLVALSGHTFGKNQCRF 180  
QY 181 IMDRLYNFSNTGLPDPPTLNTTYLQTLRGLCPLNGNLSALVDFDLRTPTTFDNKYYVNL 240  
DB 181 IMDRLYNFSNTGLPDPPTLNTTYLQTLRGLCPLNGNLSALVDFDLRTPTTFDNKYYVNL 240  
QY 241 QKGLIQSDOELFSSPNATDTIPLVRSFANSTOTFFNAFVEMDRMGNTITPLTGTQGI 300  
DB 241 QKGLIQSDOELFSSPNATDTIPLVRSFANSTOTFFNAFVEMDRMGNTITPLTGTQGI 300  
QY 301 NCRVNSNS 309  
DB 301 NCRVNSNS 309

## RESULT 7

US-10-274-793-12

; Sequence 12, Application US/10274793  
; Publication No. US20030153042A1  
; GENERAL INFORMATION:  
; APPLICANT: CALIFORNIA INSTITUTE OF TECHNOLOGY;  
; APPLICANT: Frances S. ARNOLD  
; APPLICANT: Zhanglin LIN  
; TITLE OF INVENTION: Expression of Functional Eukaryotic  
; FILE REFERENCE: 3369/18834-US3  
; CURRENT FILING DATE: 2002-10-21  
; PRIOR FILING DATE: 2002-10-21  
; PRIOR APPLICATION NUMBER: US/09/654,493  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: US 09/538,591  
; PRIOR FILING DATE: 2000-03-27  
; PRIOR APPLICATION NUMBER: US 09/247,232  
; PRIOR FILING DATE: 1999-02-09  
; PRIOR APPLICATION NUMBER: US 60/094,403  
; PRIOR FILING DATE: 1998-07-28  
; PRIOR APPLICATION NUMBER: US 60/106,840  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 12

; LENGTH: 309  
; TYPE: PRT  
; ORGANISM: Escherichia coli  
US-10-274-793-12

Query Match 99.1%; Score 1594; DB 14; Length 309;  
Best Local Similarity 99.0%; Pred. No. 2.4e-161;  
Matches 306; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MQLTPTFYDNCNPNVSNIVRDTIVNELRSDPRIAASILRLHFHDCFVNGCDASILLDNTT 60  
DB 1 MQLTPTFYDNCNPNVSNIVRDTIVNELRSDPRIAASILRLHFHDCFVNGCDASILLDNTT 60  
QY 61 SPRTKDAFGNANSARGFPVIDRMKAAVESACPTVSCADLLTIAAQSVTLAGGFSWRV 120  
DB 61 SPRTKDAFGNANSARGFPVIDRMKAAVESACPTVSCADLLTIAAQSVTLAGGFSWRV 120

QY 121 PLGRDLSLOAFDLANANLPAPFPTLPOLKDSFRNVLNRSDDLVALSGHTFGKNQCRF 180  
DB 121 PLGRDLSLOAFDLANANLPAPFPTLPOLKDSFRNVLNRSDDLVALSGHTFGKNQCRF 180  
QY 181 IMDRLYNFSNTGLPDPPTLNTTYLQTLRGLCPLNGNLSALVDFDLRTPTTFDNKYYVNL 240  
DB 181 IMDRLYNFSNTGLPDPPTLNTTYLQTLRGLCPLNGNLSALVDFDLRTPTTFDNKYYVNL 240  
QY 241 QKGLIQSDOELFSSPNATDTIPLVRSFANSTOTFFNAFVEMDRMGNTITPLTGTQGI 300  
DB 241 QKGLIQSDOELFSSPNATDTIPLVRSFANSTOTFFNAFVEMDRMGNTITPLTGTQGI 300  
QY 301 NCRVNSNS 309  
DB 301 NCRVNSNS 309

## RESULT 8

US-10-274-793-14

; Sequence 14, Application US/10274793  
; Publication No. US20030153042A1  
; GENERAL INFORMATION:  
; APPLICANT: CALIFORNIA INSTITUTE OF TECHNOLOGY;  
; APPLICANT: Frances S. ARNOLD  
; APPLICANT: Zhanglin LIN  
; TITLE OF INVENTION: Expression of Functional Eukaryotic  
; FILE REFERENCE: 3369/18834-US3  
; CURRENT APPLICATION NUMBER: US/10/274,793  
; CURRENT FILING DATE: 2002-10-21  
; PRIOR APPLICATION NUMBER: US/09/654,493  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: US 09/538,591  
; PRIOR FILING DATE: 2000-03-27  
; PRIOR APPLICATION NUMBER: US 09/247,232  
; PRIOR FILING DATE: 1999-02-09  
; PRIOR APPLICATION NUMBER: US 60/094,403  
; PRIOR FILING DATE: 1998-07-28  
; PRIOR APPLICATION NUMBER: US 60/106,840  
; PRIOR FILING DATE: 1998-11-03  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 14  
; LENGTH: 309  
; TYPE: PRT  
; ORGANISM: Escherichia coli  
US-10-274-793-14

Query Match 98.9%; Score 1590; DB 14; Length 309;  
Best Local Similarity 99.0%; Pred. No. 6.3e-161;  
Matches 306; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MQLTPTFYDNCNPNVSNIVRDTIVNELRSDPRIAASILRLHFHDCFVNGCDASILLDNTT 60  
DB 1 MQLTPTFYDNCNPNVSNIVRDTIVNELRSDPRIAASILRLHFHDCFVNGCDASILLDNTT 60  
QY 61 SPRTKDAFGNANSARGFPVIDRMKAAVESACPTVSCADLLTIAAQSVTLAGGFSWRV 120  
DB 61 SPRTKDAFGNANSARGFPVIDRMKAAVESACPTVSCADLLTIAAQSVTLAGGFSWRV 120  
QY 121 PLGRDLSLOAFDLANANLPAPFPTLPOLKDSFRNVLNRSDDLVALSGHTFGKNQCRF 180  
DB 121 PLGRDLSLOAFDLANANLPAPFPTLPOLKDSFRNVLNRSDDLVALSGHTFGKNQCRF 180  
QY 181 IMDRLYNFSNTGLPDPPTLNTTYLQTLRGLCPLNGNLSALVDFDLRTPTTFDNKYYVNL 240  
DB 181 IMDRLYNFSNTGLPDPPTLNTTYLQTLRGLCPLNGNLSALVDFDLRTPTTFDNKYYVNL 240  
QY 241 QKGLIQSDOELFSSPNATDTIPLVRSFANSTOTFFNAFVEMDRMGNTITPLTGTQGI 300  
DB 241 QKGLIQSDOELFSSPNATDTIPLVRSFANSTOTFFNAFVEMDRMGNTITPLTGTQGI 300

```
QY 301 NCRVNSNS 309
Db 301 NCRVNSNS 309

RESULT 9
US-10-274-793-20
; Sequence 20, Application US/10274793
; Publication No. US20030153042A1
; GENERAL INFORMATION:
; APPLICANT: CALIFORNIA INSTITUTE OF TECHNOLOGY;
; APPLICANT: Frances S. ARNOLD
; APPLICANT: Zhanglin LIN
; TITLE OF INVENTION: Expression of Functional Eukaryotic
; TITLE OF INVENTION: proteins
; FILE REFERENCE: 3369/1E804-US3
; CURRENT APPLICATION NUMBER: US/10/274,793
; CURRENT FILING DATE: 2002-10-21
; PRIOR APPLICATION NUMBER: US/09/654,493
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: US 09/538,591
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: US 09/247,232
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: US 60/094,403
; PRIOR FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: US 60/106,840
; PRIOR FILING DATE: 1998-11-03
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 20
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-274-793-20

Query Match 98.1%; Score 1577; DB 14; Length 309;
Best Local Similarity 98.4%; Pred. No. 1.5e-159;
Matches 304; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 MQLPTFYDNCSPNVNIVRDTIVNELRSPRIAASILRLHFHDFCVNGCDASILLDNTT 60
Db 1 MQLPTFYDNCSPNVNIVRDTIVNELRSPRIAASILRLHFHDFCVNGCDASILLDNTT 60
QY 61 SFRTEKDAFGNANSARGFPVIDRMKAAVESACPRTVSCADLLIAAQQSVTLAGGFSWRV 120
Db 61 SFRTEKDAFGNANSARGFPVIDRMKAAVESACPRTVSCADLLIAAQQSVTLAGGFSWRV 120
QY 121 PLGRDLSQAFLDLANANLPAPFTTLPOLKDSFRNVLNRSDDLVALSGGHTFGKNQCRF 180
Db 121 PLGRDLSQAFLDLANANLPAPFTTLPOLKDSFRNVLNRSDDLVALSGGHTFGKNQCRF 180
QY 181 IMRLYNFNTGLPDPNTLNTYLTQLRGLCPNLGNLSALVDFDLRTPTTFDNKYVYNLEE 240
Db 181 IMRLYNFNTGLPDPNTLNTYLTQLRGLCPNLGNLSALVDFDLRTPTTFDNKYVYNLEE 240
QY 241 OKGLIQSOQLFSSPNATDTIPLVRSFANSTQTFEFNAFVEMDRMGNTIPLTGTQGI 300
Db 241 OKGLIQSOQLFSSPNATDTIPLVRSFANSTQTFEFNAFVEMDRMGNTIPLTGTQGI 300
QY 301 NCRVNSNS 309
Db 301 NCRVNSNS 309

RESULT 10
US-10-274-793-24
; Sequence 24, Application US/10274793
; Publication No. US20030153042A1
; GENERAL INFORMATION:
; APPLICANT: CALIFORNIA INSTITUTE OF TECHNOLOGY;
; APPLICANT: Frances S. ARNOLD
; APPLICANT: Zhanglin LIN
; TITLE OF INVENTION: Expression of Functional Eukaryotic
; TITLE OF INVENTION: proteins
; FILE REFERENCE: 3369/1E804-US3
; CURRENT APPLICATION NUMBER: US/10/274,793
; CURRENT FILING DATE: 2002-10-21
; PRIOR APPLICATION NUMBER: US/09/654,493
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: US 09/538,591
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: US 09/247,232
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: US 60/094,403
; PRIOR FILING DATE: 1998-07-28
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 20
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-274-793-24

Query Match 98.1%; Score 1577; DB 14; Length 309;
Best Local Similarity 98.4%; Pred. No. 1.5e-159;
Matches 304; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 MQLPTFYDNCSPNVNIVRDTIVNELRSPRIAASILRLHFHDFCVNGCDASILLDNTT 60
Db 1 MQLPTFYDNCSPNVNIVRDTIVNELRSPRIAASILRLHFHDFCVNGCDASILLDNTT 60
QY 61 SFRTEKDAFGNANSARGFPVIDRMKAAVESACPRTVSCADLLIAAQQSVTLAGGFSWRV 120
Db 61 SFRTEKDAFGNANSARGFPVIDRMKAAVESACPRTVSCADLLIAAQQSVTLAGGFSWRV 120
QY 121 PLGRDLSQAFLDLANANLPAPFTTLPOLKDSFRNVLNRSDDLVALSGGHTFGKNQCRF 180
Db 121 PLGRDLSQAFLDLANANLPAPFTTLPOLKDSFRNVLNRSDDLVALSGGHTFGKNQCRF 180
QY 181 IMRLYNFNTGLPDPNTLNTYLTQLRGLCPNLGNLSALVDFDLRTPTTFDNKYVYNLEE 240
Db 181 IMRLYNFNTGLPDPNTLNTYLTQLRGLCPNLGNLSALVDFDLRTPTTFDNKYVYNLEE 240
QY 241 OKGLIQSOQLFSSPNATDTIPLVRSFANSTQTFEFNAFVEMDRMGNTIPLTGTQGI 300
Db 241 OKGLIQSOQLFSSPNATDTIPLVRSFANSTQTFEFNAFVEMDRMGNTIPLTGTQGI 300
QY 301 NCRVNSNS 309
Db 301 NCRVNSNS 309
```

```
; TITLE OF INVENTION: Expression of Functional Eukaryotic
; TITLE OF INVENTION: proteins
; FILE REFERENCE: 3369/1E804-US3
; CURRENT APPLICATION NUMBER: US/10/274,793
; CURRENT FILING DATE: 2002-10-21
; PRIOR APPLICATION NUMBER: US/09/654,493
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: US 09/538,591
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: US 09/247,232
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: US 60/094,403
; PRIOR FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: US 60/106,840
; PRIOR FILING DATE: 1998-11-03
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 24
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-274-793-24

Query Match 98.1%; Score 1577; DB 14; Length 309;
Best Local Similarity 98.4%; Pred. No. 1.5e-159;
Matches 304; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 MQLPTFYDNCSPNVNIVRDTIVNELRSPRIAASILRLHFHDFCVNGCDASILLDNTT 60
Db 1 MQLPTFYDNCSPNVNIVRDTIVNELRSPRIAASILRLHFHDFCVNGCDASILLDNTT 60
QY 61 SFRTEKDAFGNANSARGFPVIDRMKAAVESACPRTVSCADLLIAAQQSVTLAGGFSWRV 120
Db 61 SFRTEKDAFGNANSARGFPVIDRMKAAVESACPRTVSCADLLIAAQQSVTLAGGFSWRV 120
QY 121 PLGRDLSQAFLDLANANLPAPFTTLPOLKDSFRNVLNRSDDLVALSGGHTFGKNQCRF 180
Db 121 PLGRDLSQAFLDLANANLPAPFTTLPOLKDSFRNVLNRSDDLVALSGGHTFGKNQCRF 180
QY 181 IMRLYNFNTGLPDPNTLNTYLTQLRGLCPNLGNLSALVDFDLRTPTTFDNKYVYNLEE 240
Db 181 IMRLYNFNTGLPDPNTLNTYLTQLRGLCPNLGNLSALVDFDLRTPTTFDNKYVYNLEE 240
QY 241 OKGLIQSOQLFSSPNATDTIPLVRSFANSTQTFEFNAFVEMDRMGNTIPLTGTQGI 300
Db 241 OKGLIQSOQLFSSPNATDTIPLVRSFANSTQTFEFNAFVEMDRMGNTIPLTGTQGI 300
QY 301 NCRVNSNS 309
Db 301 NCRVNSNS 309

RESULT 11
US-10-274-793-22
; Sequence 22, Application US/10274793
; Publication No. US20030153042A1
; GENERAL INFORMATION:
; APPLICANT: CALIFORNIA INSTITUTE OF TECHNOLOGY;
; APPLICANT: Frances S. ARNOLD
; APPLICANT: Zhanglin LIN
; TITLE OF INVENTION: Expression of Functional Eukaryotic
; TITLE OF INVENTION: proteins
; FILE REFERENCE: 3369/1E804-US3
; CURRENT APPLICATION NUMBER: US/10/274,793
; CURRENT FILING DATE: 2002-10-21
; PRIOR APPLICATION NUMBER: US/09/654,493
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: US 09/538,591
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: US 09/247,232
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: US 60/094,403
; PRIOR FILING DATE: 1998-07-28
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 24
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-274-793-22
```

; PRIOR APPLICATION NUMBER: US 60/106,840  
; PRIOR FILING DATE: 1998-11-03  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 22  
; LENGTH: 309  
; TYPE: PRT  
; ORGANISM: Escherichia coli  
US-10-274-793-22

Query Match 97.8%; Score 1572; DB 14; Length 309;  
Best Local Similarity 93.1%; Pred. No. 5.3e-159;  
Matches 303; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 MQLPTFYDNCSPVNSIVRDTIIVNELRSDPPIAASILRLHFDHCFVNGCDASILLDNTT 60  
DB 1 MQLPTFYDNCSPVNSIVRDTIIVNELRSDPPIAASILRLHFDHCFVNGCDASILLDNTT 60  
QY 61 SPTEKDAFGNANSARGFPVIDRMKAIVESACPTVSCADLLTIAAQSVTLAGGSPWRV 120  
DB 61 SPTEKDAFGNANSARGFPVIDRMKAIVESACPTVSCADLLTIAAQSVTLAGGSPWRV 120  
QY 121 PGGRRDSLOAFDLANANLPAPFTLPOLKDSFRNVLNRSSDLVALSGHTEGKNOCRF 180  
DB 121 PGGRRDSLOAFDLANANLPAPFTLPOLKDSFRNVLNRSSDLVALSGHTEGKNOCRF 180  
QY 181 IMRLYNFNTGLPDPPTLNTTYLQTLRGLCPNGNLSALVDFDLRTPTIFDNKYVNL 240  
DB 181 IMRLYNFNTGLPDPPTLNTTYLQTLRGLCPNGNLSALVDFDLRTPTIFDNKYVNL 240  
QY 241 QKGLIQSDQELFSSPDATDITPLVRFSANSTQTFNFAVEAMDRMGNTITLTGTQGOIRL 300  
DB 241 QKGLIQSDQELFSSPDATDITPLVRFSANSTQTFNFAVEAMDRMGNTITLTGTQGOIRL 300  
QY 301 NCRVNSNS 309  
DB 301 NCRVNSNS 309

RESULT 12  
US-10-424-595-278555  
; Sequence 278555, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 278555  
; LENGTH: 352  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_93558C.1.pap  
US-10-424-595-278555

Query Match 57.3%; Score 921; DB 12; Length 352;  
Best Local Similarity 59.1%; Pred. No. 2.2e-89;  
Matches 182; Conservative 40; Mismatches 84; Indels 2; Gaps 2;

QY 2 QLTPTFYDNCSPVNSIVRDTIIVNELRSDPPIAASILRLHFDHCFVNGCDASILLDNTTS 61  
DB 27 QLTPTFYRETCPNLPFIVFGIFDASFTDPRIGASILRLHFDHCFVNGCDASILLDNTT 86  
QY 62 FTEKDAFGNANSARGFPVIDRMKAIVESACPTVSCADLLTIAAQSVTLAGGSPWRV 121  
DB 87 ISEQDALPNNISIRGLDVVNDIKTAVKACPGWSCADILTLASEITSSVLGGGPDWKVP 146

QY 122 LGRDLSLOAFDLANANLPAPFTLPOLKDSFRNVLNRSSDLVALSGHTEGKNOCRF 181  
DB 147 LGRDLSLTANRTLANQNLPAPFENLTQKASFAVOGLN-TLDLVTLSGGHTEGKNOCRF 205  
QY 182 MDRLYNFNTGLPDPPTLNTTYLQTLRGLCPNGNLSALVDFDLRTPTIFDNKYVNL 241  
DB 206 INRLYNFNTGNPDPTLNTTYLQTLRGLCPNGNLSALVDFDLRTPTIFDNKYVNL 265  
QY 242 KGLIQSDQELFSSPDATDITPLVRFSANSTQTFNFAVEAMDRMGNTITLTGTQGOIRL 301  
DB 266 NGLQSDQELFSTPGA-DTIPVNSFSSMONTFFSNFRSMIKMGNTIGVTGDEGEIRLQ 324  
QY 302 CRVNSNS 309  
DB 325 CNFVNGDS 332

RESULT 13  
US-10-424-599-250229  
; Sequence 250229, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 250229  
; LENGTH: 354  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_67986C.1.pap  
US-10-424-599-250229

Query Match 56.7%; Score 912.5; DB 12; Length 354;  
Best Local Similarity 57.5%; Pred. No. 1.8e-88;  
Matches 177; Conservative 50; Mismatches 78; Indels 3; Gaps 3;

QY 2 QLTPTFYDNCSPVNSIVRDTIIVNELRSDPPIAASILRLHFDHCFVNGCDASILLDNTTS 61  
DB 28 QLDPSFYRTCPRVHSIVREVNRVSKXPRMLASILRLHFDHCFVNGCDASILLDNTAT 87  
QY 62 FTEKDAFGNANSARGFPVIDRMKAIVESACPTVSCADLLTIAAQSVTLAGGSPWRV 121  
DB 88 ISEQDALPNNISIRGLDVVNDIKTAVKACPGWSCADILTLASEITSSVLGGGPDWKVP 147  
QY 122 LGRDLSLOAFDLANANLPAPFTLPOLKDSFRNVLNRSSDLVALSGHTEGKNOCRF 181  
DB 148 LGRDLSLTANRTLANQNLPAPFENLTQKASFAVOGLD-TTDLVLSGAHTFGRAHCNFI 206  
QY 182 MDRLYNFNTGLPDPPTLNTTYLQTLRGLCPNGNLSALVDFDLRTPTIFDNKYVNL 241  
DB 207 LDRLYNFNTGTXFDPTLDTTYLQQLRQICP-NGGPNLNVFDPVTPDKIDRVVFSNLQVK 265  
QY 242 KGLIQSDQELFSSPDATDITPLVRFSANSTQTFNFAVEAMDRMGNTITLTGTQGOIRL 301  
DB 266 KGLQSDQELFSTPGA-DTIPVNSFSSDQKVFVDFDAFEASIMKMGNTIGVTGKKEIRKH 324  
QY 302 CRVNSNS 309  
DB 325 CNFVAKKS 332

RESULT 14  
US-10-424-599-251628  
; Sequence 251628, Application US/10424599



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 6, 2004, 18:49:16 ; Search time 11.1649 Seconds  
(without alignments)  
1428.803 Million cell updates/sec

Title: US-09-246-451A-17

Perfect score: 1608

Sequence: 1 MQLTPTFYDNCSPVNSIVR.....PLGTGQIRINCRVNSNS 309

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

1: /cgn2\_6/ptodata/2/iaa/5A COMB.pap:\*

2: /cgn2\_6/ptodata/2/iaa/5B COMB.pap:\*

3: /cgn2\_6/ptodata/2/iaa/6A COMB.pap:\*

4: /cgn2\_6/ptodata/2/iaa/6B COMB.pap:\*

5: /cgn2\_6/ptodata/2/iaa/PTUS COMB.pap:\*

6: /cgn2\_6/ptodata/2/iaa/backfiles.pap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1462	90.9	349	1	US-08-190-029A-10
2	1462	90.9	349	2	US-08-462-695-10
3	921	57.3	352	3	US-09-207-914-22
4	781.5	48.6	324	3	US-09-365-150-2
5	768.5	47.8	333	4	US-09-097-319A-2
6	727.5	45.2	318	4	US-09-615-192A-389
7	722.5	44.9	323	4	US-09-615-192A-395
8	719.5	44.7	315	4	US-09-615-192A-378
9	646.5	40.2	287	4	US-09-615-192A-387
10	594.5	37.0	324	2	US-08-671-320-11
11	594.5	37.0	324	2	US-08-368-577-11
12	594.5	37.0	324	4	US-09-207-914-11
13	591.5	36.8	324	2	US-08-671-320-13
14	591.5	36.8	324	2	US-08-668-577-13
15	591.5	36.8	324	4	US-09-207-914-13
16	590	36.7	266	4	US-09-615-192A-394
17	587.5	36.5	351	4	US-09-615-192A-397
18	566	35.2	308	4	US-09-615-192A-381
19	566	35.2	313	2	US-08-671-320-15
20	566	35.2	313	2	US-08-668-577-15
21	566	35.2	313	4	US-09-207-914-15
22	565	35.1	313	4	US-08-671-320-17
23	565	35.1	313	2	US-08-668-577-17
24	565	35.1	313	4	US-09-207-914-17
25	503	31.3	223	4	US-09-615-192A-396
26	458.5	28.5	201	4	US-09-615-192A-392
27	433.5	27.0	208	4	US-09-615-192A-385

28 372.5 23.2 138 4 US-09-615-192A-384 Sequence 384, App  
29 365.5 22.7 161 4 US-09-615-192A-388 Sequence 388, App  
30 343 21.3 143 4 US-09-615-192A-401 Sequence 401, App  
31 331 20.6 157 4 US-09-615-192A-399 Sequence 399, App  
32 331 20.6 170 4 US-09-615-192A-332 Sequence 332, App  
33 312.5 19.4 179 4 US-09-615-192A-376 Sequence 376, App  
34 298.5 18.6 121 1 US-08-190-029A-12 Sequence 12, Appl  
35 298.5 18.6 121 2 US-08-462-695-12 Sequence 12, Appl  
36 288 17.9 202 4 US-09-615-192A-386 Sequence 386, App  
37 278.5 17.3 120 4 US-09-615-192A-393 Sequence 393, App  
38 268.5 16.7 117 4 US-09-615-192A-400 Sequence 400, App  
39 244.5 15.2 114 4 US-09-615-192A-347 Sequence 347, App  
40 241 15.0 111 4 US-09-615-192A-379 Sequence 379, App  
41 237.5 14.8 95 4 US-09-615-192A-390 Sequence 390, App  
42 235.5 14.6 118 4 US-09-615-192A-333 Sequence 333, App  
43 200 12.4 120 4 US-09-615-192A-392 Sequence 392, App  
44 187 11.6 103 4 US-09-615-192A-398 Sequence 398, App  
45 121.5 7.6 365 4 US-09-748-264A-2 Sequence 2, Appl

## ALIGNMENTS

RESULT 1  
US-08-190-029A-10  
; Sequence 10, Application US/08190029A  
; Patent No. 5736363  
; GENERAL INFORMATION:  
; APPLICANT: EDWARDS, Richard Mark  
; APPLICANT: BARDEN, Lindsey  
; TITLE OF INVENTION: IGF-II ANALOGUES  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: ALLEGRETTI & WITCOFF, LTD.  
; STREET: 10 S. WACKER DRIVE, SUITE 3000  
; CITY: CHICAGO  
; STATE: ILLINOIS  
; COUNTRY: U.S.A.  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/190,029A  
; FILING DATE: 28-FEB-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/GB92/01389  
; FILING DATE: 27-JUL-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9202401.7  
; FILING DATE: 05-FEB-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9116325.3  
; FILING DATE: 29-JUL-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: JOHN J. MCORNNELL  
; REGISTRATION NUMBER: 26,949  
; REFERENCE/DOCKET NUMBER: 94,062  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-715-1000  
; TELEFAX: 312-715-1234  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 349 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-190-029A-10

Query March 90.9%; Score 1462; DB 1; Length 349;

MOLECULE TYPE: protein  
US-08-462-695-10  
Query Match 90.9%; Score 1462; DB 2; Length 349;  
Best Local Similarity 99.6%; Pred. No. 3.8e-151;  
Matches 281; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MQLTFTFYDNCSPVNSIVRDTIVNELRSDPRTIAASILRLHFDHCFVNGCDASILLDNTT 60  
Db 1 MQLTFTFYDNCSPVNSIVRDTIVNELRSDPRTIAASILRLHFDHCFVNGCDASILLDNTT 60  
QY 61 SPTKDAFGNANSARGFPVIDRMKAIVESACPTVSCADLLTIAAQSVTLAGGSPWRV 120  
Db 61 SPTKDAFGNANSARGFPVIDRMKAIVESACPTVSCADLLTIAAQSVTLAGGSPWRV 120  
QY 121 PLGRDLSLOAFDLANANLPAPFTTLPOLKDSFRNVGLNRSSDLVALSGHTEGKNCQRF 180  
Db 121 PLGRDLSLOAFDLANANLPAPFTTLPOLKDSFRNVGLNRSSDLVALSGHTEGKNCQRF 180  
QY 181 IMRLYNFSNTGLPDPTINTTYLQTLRGLCPNLGNLSALVDFDLRTPTTFDNKYVNL 240  
Db 181 IMRLYNFSNTGLPDPTINTTYLQTLRGLCPNLGNLSALVDFDLRTPTTFDNKYVNL 240  
QY 241 QKGLIQSDQELFSSPDATDTIPLVRSFANSTQTFNFAFVAM 282  
Db 241 QKGLIQSDQELFSSPDATDTIPLVRSFANSTQTFNFAFVAM 282

RESULT 3  
US-09-207-914-22  
; Sequence 22, Application US/09207914A  
; Patent No. 6586583  
; GENERAL INFORMATION:  
; APPLICANT: Vierling Jr., Richard A.  
; TITLE OF INVENTION: A Soybean Peroxidase Gene Family and an Assay for  
; TITLE OF INVENTION: Detecting Soybean Peroxidase Activity  
; FILE REFERENCE: Soybean Peroxidase Gene Family  
; CURRENT APPLICATION NUMBER: US/09/207,914A  
; CURRENT FILING DATE: 1998-12-09  
; EARLIER APPLICATION NUMBER: US 08/868,577  
; EARLIER FILING DATE: 1997-06-04  
; EARLIER APPLICATION NUMBER: US 08/671,320  
; EARLIER FILING DATE: 1995-10-27  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: Patent in Ver. 2.0  
; SEQ ID NO 22  
; LENGTH: 352  
; TYPE: PRT  
; ORGANISM: Glycine max  
US-09-207-914-22

Query Match 57.3%; Score 921; DB 4; Length 352;  
Best Local Similarity 59.1%; Pred. No. 5e-92;  
Matches 182; Conservative 40; Mismatches 84; Indels 2; Gaps 2;  
QY 2 QLTFTFYDNCSPVNSIVRDTIVNELRSDPRTIAASILRLHFDHCFVNGCDASILLDNTT 61  
Db 27 QLTFTFYRETCPNLFPIVGFIDASFDPPIGRISIMRLHFDHCFVQCGDGSVILNNTDT 86  
QY 62 PRTKDAFGNANSARGFPVIDRMKAIVESACPTVSCADLLTIAAQSVTLAGGSPWRV 121  
Db 87 TESEODALPNINSIGLDVNDIKTAVENSCPTVSCADILIAAAEIASVLGGGPGWVPV 146  
QY 122 LGRDLSLOAFDLANANLPAPFTTLPOLKDSFRNVGLNRSSDLVALSGHTEGKNCQRF 181  
Db 147 LGRDLSLTANRTLANONLPAPFENLTQLKASPAVQGLN-TLQVLTLSGHTFGRACSTF 205  
QY 162 MRLYNFSNTGLPDPTINTTYLQTLRGLCPNLGNLSALVDFDLRTPTTFDNKYVNL 241  
Db 206 INRLYNFSNTGNPDPTINTTYLQTLRGLCPNLGNLSALVDFDLRTPTTFDNKYVNL 255  
QY 242 KGLIQSDQELFSSPDATDTIPLVRSFANSTQTFNFAFVAMDMRMGNITLTGTGQIRLN 301

Best Local Similarity 99.6%; Pred. No. 3.8e-151;  
Matches 281; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MQLTFTFYDNCSPVNSIVRDTIVNELRSDPRTIAASILRLHFDHCFVNGCDASILLDNTT 60  
Db 1 MQLTFTFYDNCSPVNSIVRDTIVNELRSDPRTIAASILRLHFDHCFVNGCDASILLDNTT 60  
QY 61 SPTKDAFGNANSARGFPVIDRMKAIVESACPTVSCADLLTIAAQSVTLAGGSPWRV 120  
Db 61 SPTKDAFGNANSARGFPVIDRMKAIVESACPTVSCADLLTIAAQSVTLAGGSPWRV 120  
QY 121 PLGRDLSLOAFDLANANLPAPFTTLPOLKDSFRNVGLNRSSDLVALSGHTEGKNCQRF 180  
Db 121 PLGRDLSLOAFDLANANLPAPFTTLPOLKDSFRNVGLNRSSDLVALSGHTEGKNCQRF 180  
QY 181 IMRLYNFSNTGLPDPTINTTYLQTLRGLCPNLGNLSALVDFDLRTPTTFDNKYVNL 240  
Db 181 IMRLYNFSNTGLPDPTINTTYLQTLRGLCPNLGNLSALVDFDLRTPTTFDNKYVNL 240  
QY 241 QKGLIQSDQELFSSPDATDTIPLVRSFANSTQTFNFAFVAM 282  
Db 241 QKGLIQSDQELFSSPDATDTIPLVRSFANSTQTFNFAFVAM 282

RESULT 2  
US-08-462-695-10  
; Sequence 10, Application US/08462695  
; Patent No. 5854025  
; GENERAL INFORMATION:  
; APPLICANT: EDWARDS, Richard Mark  
; APPLICANT: BAWDEN, Lindsey  
; TITLE OF INVENTION: IGF-II ANALOGUES  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: BANNER & ALLEGRETTI, LTD.  
; STREET: 10 S. WACKER DRIVE, SUITE 3000  
; CITY: CHICAGO  
; STATE: ILLINOIS  
; COUNTRY: U.S.A.  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/462,695  
; FILING DATE: 5-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/190,029  
; FILING DATE: 28-FEB-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/GB92/01389  
; FILING DATE: 27-JUL-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9202401.7  
; FILING DATE: 05-FEB-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9116325.3  
; FILING DATE: 29-JUL-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: JOHN J. McDONNELL  
; REGISTRATION NUMBER: 26,949  
; REFERENCE/DOCKET NUMBER: 94,062-A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-715-1000  
; TELEFAX: 312-715-1234  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 349 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear

Db 266 NGLQSDQELFSTPGA-DTPIVNSFSSNQNTFFSNRVSIMKMGNIGVLTGDEGEIRLQ 324  
QY 302 CRVANSNS 309  
Db 325 CNFVNGDS 332

## RESULT 4

US-09-365-150-2  
; Sequence 2, Application US/09365150  
; Patent No. 6278041  
; GENERAL INFORMATION:  
; APPLICANT: Lagrimini, Mark  
; APPLICANT: Desai, Nalini  
; TITLE OF INVENTION: No. 6278041el Peroxidase Gene Sequences  
; CURRENT APPLICATION NUMBER: JS/09/365,150  
; CURRENT FILING DATE: 1999-07-30  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 324  
; TYPE: PRT  
; ORGANISM: Kocotiana tomentosiformis  
US-09-365-150-2

Query Match 48.6%; Score 781.5; DB 3; Length 324;  
Best Local Similarity 50.3%; Pred. No. 7.7e-77;  
Matches 154; Conservative 59; Mismatches 88; Indels 5; Gaps 5;  
QY 2 QLTPTFYNSCNVSNIVRDTIVNELRSDPRIAASILRLHFHDCFVNGCDASILLDTTSF 61  
Db 23 QLSATFYDSTCPNVTSTVRGVMQDQRTDARAGAKIIRLHFHDCFVNGCDGSILLD-IDG 81  
QY 62 RTEKDAFNGANSARGFPVIDRMKAAVESACPTVSCADLLTIAAQSVTLAGGSPWRVP 121  
Db 82 TQTEKDAAPNV-GAGGDFIVDDIKTALENVCPGWSCADILSLASIGVALAEGPSQWL 140  
QY 122 LGRDLSQAFDLANAKLPAPFTTLPQKDSFRNVLNRSDDLVALSGGHTGKNCRTI 181  
Db 141 FGKNSLTNRSEANS-ESPETPAVMTPLTNKGM-DTLNLAGSGAHTGRACGTF 199  
QY 182 MDRLNFSNTGLPDTLNTTY-QTLRGLCPNLGNL-SALVDFDLRPTPTTFIDNKYYVNL 240  
Db 200 EQRLENFSGNPDPTVDATFLQTLQICPGQNGNNTFTNLIDSTPNDFDNDYFNLQN 259  
QY 241 QKGLTQSDQLFSSPDATDTIPLVRSFANSTQTFENAFVEMDRMGNITPLTGTQCI 300  
Db 260 NQSLQTDQELFST-SGSATIAVRYAGSQTFDDFVSSMIKLGNSPLTGTNGEIRT 318  
QY 301 NCRVYN 306  
Db 319 DCRVYN 324

## RESULT 5

US-09-097-319A-2  
; Sequence 2, Application US/09097319A  
; Patent No. 6384207  
; GENERAL INFORMATION:  
; APPLICANT: Ainley, Michael  
; APPLICANT: Armstrong, Katherine  
; APPLICANT: Belmar, Scott  
; APPLICANT: Folkerts, Otto  
; APPLICANT: Hopkins, Nicole  
; APPLICANT: Menke, Michael A.  
; APPLICANT: Paredy, Dayakar  
; APPLICANT: Petolino, Joseph F.  
; APPLICANT: Smith, Kelley  
; APPLICANT: Woosley, Aaron  
; TITLE OF INVENTION: Regulatory Sequences for Transgenic Plants  
; NUMBER OF SEQUENCES: 59  
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: DowElanco Patent Department  
; STREET: 9330 Zionsville Road  
; CITY: Indianapolis  
; STATE: Indiana  
; COUNTRY: USA  
; ZIP: 46268  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/097,319A  
; FILING DATE:  
; CLASSIFICATION: 800  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Stuart, Donald R  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 317 337 4816  
; TELEFAX: 317 337 4847  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 333 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-097-319A-2

Query Match 47.8%; Score 768.5; DB 4; Length 333;  
Best Local Similarity 49.7%; Pred. No. 2.1e-75;  
Matches 152; Conservative 52; Mismatches 99; Indels 3; Gaps 2;  
QY 3 LPTFTVNSCNVSNIVRDTIVNELRSDPRIAASILRLHFHDCFVNGCDASILLDTTSF 62  
Db 30 LFPQFYDHSCKPAXE-VQSIQAQVAKETRMASLRLHFHDCFVNGCDASVLLDNSSI 89  
QY 63 RTEKDAFNGANSARGFPVIDRMKAAVESACPTVSCADLLTIAAQSVTLAGGSPWRVP 122  
Db 90 VSEKGSNPNSILRGVEVIDQIKALEACPTVSCADIVALLAARSTALVGPPYWDVPL 149  
QY 123 GRDLSQAFDLANAKLPAPFTTLPQKDSFRNVLNRSDDLVALSGGHTGKNCQCFIM 182  
Db 150 GRDLSGASIQSNDIPAPNNTLPTIITKFKQGLN-VVDVVALSGGHTGMSRCTSF 208  
QY 183 DRLNFSNTGLPDTLNTTY-LQTLRGLCPNLGNL-SALVDFDLRPTPTTFIDNKYYVNL 242  
Db 209 QRLNYGTGNGMADSTLDVSTAANKLRQCPRSGDNNLFLDLFITPAKFDNFYKLNLAG 268  
QY 243 GLTQSDQLFSSPDATDTIPLVRSFANSTQTFENAFVEMDRMGNITPLTGTQCI 302  
Db 269 GLLSDEILLTK--SAETAALVKAYADYNLFFQHPAQSMVNGNISPLTGSQGEIRKNC 326  
QY 303 RVVNSN 308  
Db 327 RLINND 332

## RESULT 6

US-09-615-192A-389  
; Sequence 389, Application US/096-5192A  
; Patent No. 6410718  
; GENERAL INFORMATION:  
; APPLICANT: Bloksberg, Leonard N.  
; APPLICANT: Havukkala, Ilkka  
; TITLE OF INVENTION: Materials and Methods for the  
; MODIFICATION OF PLANT LIGNIN CONTENT  
; FILE REFERENCE: 11000.1003C4U  
; CURRENT APPLICATION NUMBER: US/09/615,192A  
; CURRENT FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 08/975,316  
; PRIOR FILING DATE: 1997-11-21  
; PRIOR APPLICATION NUMBER: US 08/713,000  
; PRIOR FILING DATE: 1996-09-11



```

; PRIOR APPLICATION NUMBER: US 09/169,789
; PRIOR FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 405
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 389
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Pinus radiata
US-09-615-192A-389

Query Match      45.2%; Score 727.5; DB 4; Length 318;
Best Local Similarity 49.5%; Pred. No. 5.9e-71;
Matches 151; Conservative 48; Mismatches 95; Indels 11; Gaps 4;

QY 2 QLTPTFYDNCNPNVSDIVNELRSDPRAASILRLHFDHCFVNGCDASILLDNNTS 61
DB 25 QLSSTFYAKSCPRLSIVKAVKQVAVKXKRGASIVRLHFDHCFVNGCDGSLILLDNAT 84
QY 62 FRTEKDAFGNANSARGFPVIDRMKAAVESACPTVSCADLLTIAAQSVTLAGGFSWRVP 121
DB 85 FTGEXTAGNANSARGFEVIDSIKTQVEAAACGWSVCADILITIAARDSIVELQGPWTVM 144
QY 122 LGRDLSQAFDLANANLPAPFTTLPQKDSFRNVGNRSSDLVALSGGHTFGKNCQCFI 181
DB 145 LGRDSTTASAAANNIPSPASSLSTLISSFQAEGLS-TKDLVALSGAHTTQGRCAFF 203
QY 182 MDRLYNFSNTGLPDTLNTTYLTQTLRGLCPLNGLSALVDFDLRPTTFIDNKYYVLEBQ 241
DB 204 RTRIYNEN-----INNAFATSVKANCPSAGGSNLSPLDAVTSITFDNKYYNLIKQ 256
QY 242 KGLISDQELFSSPDATDIPIVRSFANSTQTFNFAFVEAMDRMGNITPLTGTQOIRN 301
DB 257 KGLHSDQQLFNG-GSTDS--QVTYSSNQNSFFIDFTAAVYKMGNSPLTGTGQIRKN 313
QY 302 CRVUN 306
DB 314 CRKSN 318

RESULT 7
US-09-615-192A-395
; Sequence 395, Application US/096:15192A
; Patent No. 6410718
; GENERAL INFORMATION:
; APPLICANT: Bloksberg, Leonard N.
; APPLICANT: Havukkala, Ilkka
; TITLE OF INVENTION: Materials and Methods for the
; FILE REFERENCE: 11000.1003C4U
; CURRENT FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 09/615,192A
; PRIOR FILING DATE: 1997-11-21
; PRIOR FILING DATE: 1998-10-03
; PRIOR FILING DATE: 1998-10-03
; NUMBER OF SEQ ID NOS: 405
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 395
; LENGTH: 323
; TYPE: PRT
; ORGANISM: Pinus radiata
US-09-615-192A-395

Query Match      44.9%; Score 722.5; DB 4; Length 323;
Best Local Similarity 49.0%; Pred. No. 2.1e-70;
Matches 152; Conservative 49; Mismatches 90; Indels 19; Gaps 6;

QY 2 QLTPTFYDNCNPNVSDIVNELRSDPRAASILRLHFDHCFVNGCDASILLDNNTS 61
DB 28 QLSSTFYKSCPTALSIVKAVKQVAVKXKRGASILRLHFDHCFVNGCDGSLILLDSST 87

; PRIOR APPLICATION NUMBER: US 09/169,789
; PRIOR FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 405
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 389
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Pinus radiata
US-09-615-192A-389

Query Match      45.2%; Score 727.5; DB 4; Length 318;
Best Local Similarity 49.5%; Pred. No. 5.9e-71;
Matches 151; Conservative 48; Mismatches 95; Indels 11; Gaps 4;

QY 2 QLTPTFYDNCNPNVSDIVNELRSDPRAASILRLHFDHCFVNGCDASILLDNNTS 61
DB 25 QLSSTFYAKSCPRLSIVKAVKQVAVKXKRGASIVRLHFDHCFVNGCDGSLILLDNAT 84
QY 62 FRTEKDAFGNANSARGFPVIDRMKAAVESACPTVSCADLLTIAAQSVTLAGGFSWRVP 121
DB 85 FTGEXTAGNANSARGFEVIDSIKTQVEAAACGWSVCADILITIAARDSIVELQGPWTVM 144
QY 122 LGRDLSQAFDLANANLPAPFTTLPQKDSFRNVGNRSSDLVALSGGHTFGKNCQCFI 181
DB 145 LGRDSTTASAAANNIPSPASSLSTLISSFQAEGLS-TKDLVALSGAHTTQGRCAFF 203
QY 182 MDRLYNFSNTGLPDTLNTTYLTQTLRGLCPLNGLSALVDFDLRPTTFIDNKYYVLEBQ 241
DB 204 RTRIYNEN-----INNAFATSVKANCPSAGGSNLSPLDAVTSITFDNKYYNLIKQ 256
QY 242 KGLISDQELFSSPDATDIPIVRSFANSTQTFNFAFVEAMDRMGNITPLTGTQOIRN 301
DB 257 KGLHSDQQLFNG-GSTDS--QVTYSSNQNSFFIDFTAAVYKMGNSPLTGTGQIRKN 313
QY 302 CRVUN 306
DB 314 CRKSN 318

RESULT 8
US-09-615-192A-378
; Sequence 378, Application US/096:15192A
; Patent No. 6410718
; GENERAL INFORMATION:
; APPLICANT: Bloksberg, Leonard N.
; APPLICANT: Havukkala, Ilkka
; TITLE OF INVENTION: Materials and Methods for the
; FILE REFERENCE: 11000.1003C4U
; CURRENT FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 09/615,192A
; PRIOR FILING DATE: 1997-11-21
; PRIOR FILING DATE: 1998-10-03
; PRIOR FILING DATE: 1998-10-03
; NUMBER OF SEQ ID NOS: 405
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 378
; LENGTH: 315
; TYPE: PRT
; ORGANISM: Eucalyptus grandis
US-09-615-192A-378

Query Match      44.7%; Score 719.5; DB 4; Length 315;
Best Local Similarity 49.2%; Pred. No. 4.4e-70;
Matches 151; Conservative 46; Mismatches 95; Indels 15; Gaps 4;

QY 2 QLTPTFYDNCNPNVSDIVNELRSDPRAASILRLHFDHCFVNGCDASILLDNNTS 61
DB 22 KLSPSHYOSTCPKALSIVRAGVAKAINEKTRTGASLLRLHFDHCFVNGCDASILLDDTFS 81
QY 62 FRTEKDAFGNANSARGFPVIDRMKAAVESACPTVSCADLLTIAAQSVTLAGGFSWRVP 121
DB 82 FVGEXTAAPNNNSVGFVVDRIKASLEKCEGCVVSCADIVALAARDSVVLHGGFSWTVS 141
QY 122 LGRDLSQAFDLANANLPAPFTTLPQKDSFRNVGNRSSDLVALSGGHTFGKNCQCFI 181
DB 142 LGRKDSITASRSLANTSIPPPTSLSALITSPFAAQGLS-VKMMVALSGSHTTGLARCTSF 200
QY 182 MDRLYNFSNTGLPDTLNTTYLTQTLRGLCPLNGLSALVDFDLRPTTFIDNKYYVLEBQ 241
DB 201 RRRIFYDNCN-----IDTSFAHKLQKICPRIGNDSVLRQDIDTPTTFIDNLYHNLLQK 253
QY 242 KGLISDQELF--SSPDATDIPIVRSFANSTQTFNFAFVEAMDRMGNITPLTGTQOIR 299
DB 254 KGLHSDQELFNGSSVDS-----LVKAYACDTGKFFRDFAKAMKSEIKPKGNGQIR 308
QY 300 LNCRVUN 306

```

```

Db      309 KNCRKVN 315

RESULT 9
US-09-615-192A-387
; Sequence 387, Application US/09615192A
; Patent No. 6410718
; GENERAL INFORMATION:
; APPLICANT: Bloksberg, Leonard N.
; APPLICANT: Havukkala, Ilkka
; TITLE OF INVENTION: Materials and Methods for the
; TITLE OF INVENTION: Modification of Plant Lignin Content
; FILE REFERENCE: 11000.1003c4U
; CURRENT APPLICATION NUMBER: US/09/615,192A
; CURRENT FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 08/975,316
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: US 08/713,000
; PRIOR FILING DATE: 1996-09-11
; PRIOR APPLICATION NUMBER: US 09/169,789
; PRIOR FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 405
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 387
; LENGTH: 287
; TYPE: PRT
; ORGANISM: Pinus radiata
US-09-615-192A-387

Query Match      40.2%; Score 646.5; DB 4; Length 287;
Best Local Similarity 48.6%; Pred. No. 3.6e-62;
Matches 135; Conservative 46; Mismatches 78; Indels 19; Gaps 6;

QY      2 QLTPTFYDNCSPNVNIVRDITIVNELRSDPRTIAASILRLHFDHCFVNGCDASILLNDTTS 61
Db      24 KLSITFDKCPKLSIVQGVKQAVAKKRMGASLELRHFDHCFVNGCDGSVLLDNTT 83
QY      62 FTEKDAFGNANSARGFPVIDRMKAAVESACPTVSCADLLTIAAQOSVTLAGGPSWRVP 121
Db      84 FTSEKALPNNNSARGFEVIDSIKQLENACTGVVSCADILTIARUSVWVLGGPSWKVM 143
QY      122 LGRDLSQAFELDLANANLPAPFFTLPOLKDSFRNVGLNRSSDLVALSGGHTFGKQCRFI 181
Db      144 LGRDSTASISGANNIPPTSNLTSLIFQAGGLS-TKEMVALSGGHTIGQAQCKNF 202
QY      182 MDRLYNFSNTGLPDTLNTTYLQTLRGLCP-NGNLSALVDFDLRPTTFDNKYV 236
Db      203 RAIYN-----DTNIDTYSLSRCKPSTTGSGDSNLSPL---DYTTTFVFDKNYY 252
QY      237 NLEEQGLIQSDQELFSSPDATDTIPLVRSEANSTOTF 274
Db      253 NLKSKRGLLHSDQELFNG-GSTDS--HVTKYASNQNTF 287

RESULT 10
US-08-671-320-11
; Sequence 11, Application US/08671320
; Patent No. 5840558
; GENERAL INFORMATION:
; APPLICANT: VIERLING JR, RICHARD A
; TITLE OF INVENTION: A SOYBEAN PEROXIDASE GENE FAMILY AND AN
; TITLE OF INVENTION: ASSAY FOR DETECTING SOYBEAN PEROXIDASE ACTIVITY
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: 1125 SO. 103RD STREET
; CITY: SUITE 330
; STATE: NE
; COUNTRY: US
; ZIP: 68124-1076
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible

Db      309 KNCRKVN 315

RESULT 11
US-08-868-577-11
; Sequence 11, Application US/08868577
; Patent No. 5866695
; GENERAL INFORMATION:
; APPLICANT: Vierling Jr., Richard A
; TITLE OF INVENTION: A SOYBEAN PEROXIDASE GENE FAMILY AND AN
; TITLE OF INVENTION: ASSAY FOR DETECTING SOYBEAN PEROXIDASE ACTIVITY
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rotwell, Figg, Ernst & Kurz
; STREET: 555 13th Street NW, Suite 701 East
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/868,577

```

```

? FILING DATE: 04-JUN-1997
? CLASSIFICATION: 536
? ATTORNEY/AGENT INFORMATION:
? NAME: Jondle, Robert J.
? REGISTRATION NUMBER: 33,915
? REFERENCE/DOCKET NUMBER: N1227-003
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 402-333-1550
? TELEFAX: 402-333-1510
? INFORMATION FOR SEQ ID NO: 11:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 324 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? US-08-868-577-11

```

```

Query Match 37.0%; Score 594.5; DB 2; Length 324;
Best Local Similarity 39.9%; Pred. No. 2.1e-56;
Matches 123; Conservative 63; Mismatches 115; Indels 7; Gaps 5;

```

```

QY 2 QLTPTFYDNCSPVNSIVRDTIVNELRSDPRIAASILRLHFDHCFVNGCDASILLNTTS 61
Db 22 QLQGFYANSCPAEQIVLKFVHDHINAPSLAAALRMHFDHCFVRGCDASVLLNSTN 81
QY 62 FTEKDAFGNANSARGFPVIDRMKAAVESACPRVTWSCADLLTIAAQOSVTLGGPSWRVP 121
Db 82 -QAEKNAPPNL-TVRGDFDIDRIKSLVEAECPGVWSCADILTLAARDTIIVATGGFWKVP 139
QY 122 LGRDLSQAFDLIANANLPAFFTLQPKDPSFRVNGLSRSDLVASGGHTFGKNCQCFI 181
Db 140 TGRDGSVNSLTERARNIPAPSNFTTLQTLFANQGLD-LKDIVLLSGAHTIGIAHCSSL 198
QY 182 MDRLYNESNTGLPDTLNTLYLQTLRGL--CPLNGNLSALVDPLRPTIIFDNKYVYLE 239
Db 199 SNRLNFTGKGQDQSLDSEYANLAKFKCTDLNKLNTTKIEMDPGSRKRTDLSYYSHWI 258
QY 240 EOKGLIQSDQELFSSPDATDTIPLVRSFANSTOTFFNAFVEMDRMGNTIPLTGTQOQIR 299
Db 259 KRGLEFSDAALLT--NSVTKAQIIQLLEGSVENFFAEFATSIKMGRIINVKTGTGEIR 316
QY 300 LNCRVVNS 307
Db 317 KHCAFINS 324

```

```

RESULT 12
US-09-207-914-11
? Sequence 11, Application US/09207914A
? Patent No. 6586583
? GENERAL INFORMATION:
? APPLICANT: Vierling Jr., Richard A.
? TITLE OF INVENTION: A Soybean Peroxidase Gene Family and an Assay for
? FILE REFERENCE: Detecting Soybean Peroxidase Activity
? CURRENT APPLICATION NUMBER: US/09/207,914A
? CURRENT FILING DATE: 1998-12-09
? EARLIER APPLICATION NUMBER: US 08/868,577
? EARLIER FILING DATE: 1997-06-04
? EARLIER APPLICATION NUMBER: US 08/671,320
? EARLIER FILING DATE: 1995-10-27
? NUMBER OF SEQ ID NOS: 24
? SOFTWARE: Patent In Ver. 2.0
? SEQ ID NO 11
? LENGTH: 324
? TYPE: PRT
? ORGANISM: Glycine max
US-09-207-914-11

```

```

Query Match 37.0%; Score 594.5; DB 4; Length 324;
Best Local Similarity 39.9%; Pred. No. 2.1e-56;
Matches 123; Conservative 63; Mismatches 115; Indels 7; Gaps 5;

```

```

QY 2 QLTPTFYDNCSPVNSIVRDTIVNELRSDPRIAASILRLHFDHCFVNGCDASILLNTTS 61
Db 22 QLQGFYANSCPAEQIVLKFVHDHINAPSLAAALRMHFDHCFVRGCDASVLLNSTN 81
QY 62 FTEKDAFGNANSARGFPVIDRMKAAVESACPRVTWSCADLLTIAAQOSVTLGGPSWRVP 121
Db 82 -QAEKNAPPNL-TVRGDFDIDRIKSLVEAECPGVWSCADILTLAARDTIIVATGGFWKVP 139
QY 122 LGRDLSQAFDLIANANLPAFFTLQPKDPSFRVNGLSRSDLVASGGHTFGKNCQCFI 181
Db 140 TGRDGSVNSLTERARNIPAPSNFTTLQTLFANQGLD-LKDIVLLSGAHTIGIAHCSSL 198
QY 182 MDRLYNESNTGLPDTLNTLYLQTLRGL--CPLNGNLSALVDPLRPTIIFDNKYVYLE 239
Db 199 SNRLNFTGKGQDQSLDSEYANLAKFKCTDLNKLNTTKIEMDPGSRKRTDLSYYSHWI 258
QY 240 EOKGLIQSDQELFSSPDATDTIPLVRSFANSTOTFFNAFVEMDRMGNTIPLTGTQOQIR 299
Db 259 KRGLEFSDAALLT--NSVTKAQIIQLLEGSVENFFAEFATSIKMGRIINVKTGTGEIR 316
QY 300 LNCRVVNS 307
Db 317 KHCAFINS 324

```

```

RESULT 13
US-08-671-320-13
? Sequence 13, Application US/08671320
? Patent No. 5840558
? GENERAL INFORMATION:
? APPLICANT: Vierling Jr., Richard A.
? TITLE OF INVENTION: A Soybean Peroxidase Gene Family and an
? TITLE OF INVENTION: ASSAY FOR DETECTING SOYBEAN PEROXIDASE ACTIVITY
? NUMBER OF SEQUENCES: 17
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: 1125 SO. 103RD STREET
? CITY: OMAHA
? STATE: NE
? COUNTRY: US
? ZIP: 68124-1076
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent In Release #1.0, Version #1.30
? CURRENT APPLICATION NUMBER: US/08/671,320
? FILING DATE:
? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? NAME: JONDLE, ROBERT J.
? REGISTRATION NUMBER: 33,915
? REFERENCE/DOCKET NUMBER: 1227-001
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 402-398-9000
? TELEFAX: 402-398-9005
? INFORMATION FOR SEQ ID NO: 13:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 324 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
US-08-671-320-13

```

```

Query Match 36.8%; Score 591.5; DB 2; Length 324;
Best Local Similarity 39.9%; Pred. No. 4.4e-56;
Matches 123; Conservative 62; Mismatches 116; Indels 7; Gaps 5;

```

```

QY 2 QLTPTFYDNCSPVNSIVRDTIVNELRSDPRIAASILRLHFDHCFVNGCDASILLNTTS 61
Db 22 QLQGFYANSCPAEQIVLKFVHDHINAPSLAAALRMHFDHCFVRGCDASVLLNSTN 81

```

Qy	62	FRTEKDAFGNANSARGEPFVIDRMKAAVESACPRVTVCADLLITIAAQSVTLAGGFSWRVP	121
Db	82	-QAEKNAPPNL-TVRGFDIFDIKSLVEACPGVWSCADILTSARDTIVATGSPFWKVP	139
Qy	122	LGRRDSLQAFDLANANLPADPFLLPOLKDSFRNVGLNRSSDLWLGGHTTGKQCRFI	181
Db	140	TGRDGVISNLTEARDNTPASSNETTITQTFANQGLD-LKDLVLSGNATTIGIAHCSSL	198
Qy	182	MRLVNFSGNTGLPPTLNTLYLTQTLRGL--CPINGNLsalvFDLRTPTIFDNKYVNL	239
Db	199	SNRLFNFTGKGDQDPSLDSEVAANLKAFKCTDLNLTKIBMDPGSRKTFDLVYVSHV	258
Qy	240	EQKGLIGSDQLSPSPDATDIPIVRSFANSCITFNAPVEAMDRMGNITELTGQQR	299
Db	259	KRRGLFSDAALLT--NSVTKAQIIELLEGSVENFFASFATSMKWRINKWITGTEGEIR	316
Qy	300	LNCRWVNS	307
Db	317	KKCAFLNS	324

RESULT 14  
US-08-868-577-13  
? Sequence 13, Application US/08868577  
? Patent No. 5866695  
? GENERAL INFORMATION:  
? APPLICANT: Vierling Jr., Richard A  
? TITLE OF INVENTION: A SOYBEAN PEROXIDASE GENE FAMILY AND AN  
? TITLE OF INVENTION: ASSAY FOR DETECTING SOYBEAN PEROXIDASE ACTIVITY  
? NUMBER OF SEQUENCES: 19

CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Rothwell, Figg, Ernst & Kurz  
 STREET: 555 13th Street NW, Suite 701 East  
 CITY: Washington  
 STATE: DC  
 COUNTRY: USA  
 ZIP: 20004  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: WordPerfect 6.1  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/868.577  
 FILING DATE: 04-JUN-1997  
 CLASSIFICATION: 536  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Jendle, Robert J.  
 REGISTRATION NUMBER: 33,915  
 REFERENCE/DOCKET NUMBER: N1227-003  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 402-333-1550  
 TELEFAX: 402-333-1510  
 INFORMATION FOR SEQ ID NO: 13:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 324 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-868-577-13

Query Match 36.8%; Score 591.5; DB 2; Length 324;  
Best Local Similarity 39.9%; Pred. No. 4.4e-56;  
Matches 123; Conservative 62; Mismatches 116; Indels 7; Gaps 5

2 QLTPTFTDNCSPVNSVIVRDTIVNELRSDPRLAASLILPHFDHPVNGCDASILLDNTTS 61  
22 QLQGFYAKSPNAEQIVLKVFDHIHNPALAAALISMFHDCDFVRCGDSAVLLNSTTN 81  
62 FRTEKDAFGNANSARGFPVVIDRMKAAYESACPRFTVSCADLETTIAQQSVTLAGGSPWVP 121  
82 -QAEKNAPPNL- TVRGRDFFIDRIKSLVEAECPGVVSCADILTLGARDIVATGGPFMKVP 139

Search completed: April 6, 2004, 18:56:01  
Job time : 12.1649 secs

```

QY 122 LGRDLSQAFDLNANLPAPFFTLPOLKDSFRNVGLNRSSDLVALSGGHTFGKXQCRFI 181
D5 140 TGRDGVISNTEARDNKIPAPSSNFTTLQTLFANQGLD-LKDLVLLSGAHTTIGIARCCSSL 198
QY 182 MDRLYNESNTGLPPTLNTTYLQTLRGL--CPLAGNLSALVDFDLRTPTFDNKYYVNL 239
D5 199 SNRUFNFTGKGQDPSLDSEYAANLKAFKCTDNLKNTTKIEMDPGSRKTFDLSYYSHWI 258
QY 240 BQKGLIQSDQELFSSPDATDIPLVRSFANSTQTFPFAVEARDKMGNTPLTGTGQGR 299
D5 259 XBRGLFESDAALLT--NSVTAKQIIELEGSVENFAEFAISYKRMGRINVKVTGTEGEIR 316
QY 300 LNCRVNS 307
D5 317 KHC AFLNS 324

RESULT 15
US-09-207-914-13
; Sequence 13, Application US/09207914A
; Patent No. 6586583
; GENERAL INFORMATION:
; APPLICANT: Vierling Jr., Richard A.
; TITLE OF INVENTION: A Soybean Peroxidase Gene Family and an Assay for
; FILE REFERENCE: Detecting Soybean Peroxidase Activity
; CURRENT APPLICATION NUMBER: US/09/207,914A
; CURRENT FILING DATE: 1998-12-09
; EARLIER APPLICATION NUMBER: US 08/868,577
; EARLIER FILING DATE: 1997-06-04
; EARLIER APPLICATION NUMBER: US 08/671,320
; EARLIER FILING DATE: 1995-10-27
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.3
; SEQ ID NO 13
; LENGTH: 324
; TYPE: PRT
; ORGANISM: Glycine max
US-09-207-914-13

```

Query Match	36.8%;	Score 591.5;	DB 4;	Length 324;
Best Local Similarity	39.9%;	Pred. No. 4.4e-56;		
Matches 123:	Conservative 62;	Mismatches 116;	Indels 7;	Gaps 5

Qy	2	QLTFTFYDNGCPNVNIVROTVJIVNELRSDPRIAASILRLHFHDCFVNGCDASILDLNTTS	61
Db	22	QLQLGFYAKSPNAEQIWLKPVFVHDHIEHNPASLAAALIRMHFDHCFVRGCTASVLINSTIN	81
Qy	62	FRTEKDFGNNNSARGFPVIDRMKAAYESACPRTVSCADILLITIAAQSVTLAGGSPWRP	121
Db	82	-QAEKNAAPPNL-TVRGDFDIDRIKSLVENECPGVSCADILLTSARDTIVATGGSPFKVP	139
Qy	122	LGRDSSQAFLIDLANANLPAFFTLPOLKOSFRVUGLNRSSDLVALSGGHFTGKNQCRFI	181
Db	140	TGRDGWISNLTARNDIPAPSNFTTQTLFANQGLD-LKDLVLLSGAHTIGIAHCSL	198
Qy	182	MDRLYNFSNTGLEPPTLNTYTLQTLRGL--CPLNGNL.SALVDFDLRTPTIEDNKYYVNL	239
Db	199	SNRLFNFTGKGDDPSLDSHYAANLKAFTCDLNLNTKIEMDPGSRKTFDLSYSHVI	258
Qy	240	EQKGLIQSDQELFSSPDATTTIPLVRSPANSTOTFFNAFVAMDRMGNITPLTCTQGOIR	299
Db	259	KRRGLFESDAALLT--NSVTKAQIIBLLEGSVENFFAEFATSMKMGKRIINVKTCTGEIR	316
Qy	300	INCRVNS	307
Db	317	KHCAFLNS	324

Search completed: April 6, 2004, 18:56:01  
Job time : 12.1649 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein: - protein search, using sw model

Run on: April 6, 2004, 18:53:17 ; Search time 5.97557 Seconds

(without alignments)

988.783 Million cell updates/sec

Title: US-09-246-451A-17

Perfect score: 1608

Sequence: 1 MQLPTTFYDNCSEVNSIVR.....PLATGTCQIRLNCRVVNSNS 303

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 159197 seqs, 19121497 residues

Total number of hits satisfying chosen parameters: 159157

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents\_AA\_New.\*

- 1: /cgn2\_6/prodata/1/paa/US06\_NEW\_COMB.pep.\*
- 2: /cgn2\_6/prodata/1/paa/US06\_NEW\_COMB.pep.\*
- 3: /cgn2\_6/prodata/1/paa/US07\_NEW\_COMB.pep.\*
- 4: /cgn2\_6/prodata/1/paa/US08\_NEW\_COMB.pep.\*
- 5: /cgn2\_6/prodata/1/paa/US09\_NEW\_COMB.pep.\*
- 6: /cgn2\_6/prodata/1/paa/US10\_NEW\_COMB.pep.\*
- 7: /cgn2\_6/prodata/1/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	768.5	47.8	333	6	US-10-759-602-2
2	767	47.7	318	6	Sequence 2, Appli
3	746.5	46.4	325	6	Sequence 45022, A
4	736.5	45.8	347	6	Sequence 45822, A
5	702	43.7	314	6	Sequence 40690, A
6	654	40.7	337	6	Sequence 47189, A
7	625	38.9	357	6	Sequence 39724, A
8	611	38.0	377	6	Sequence 39724, A
9	577.5	35.9	315	6	Sequence 46170, A
10	544.5	33.9	357	6	Sequence 46411, A
11	519.5	32.3	338	6	Sequence 44615, A
12	503.5	31.3	209	6	Sequence 43455, A
13	490.5	30.5	341	6	Sequence 44200, A
14	489	30.4	251	6	Sequence 42477, A
15	468	29.1	347	6	Sequence 44634, A
16	457	28.4	263	6	Sequence 35185, A
17	452	28.1	401	6	Sequence 46344, A
18	447.5	27.8	230	6	Sequence 51301, A
19	435	27.1	203	6	Sequence 40445, A
20	421	26.2	200	6	Sequence 39630, A
21	405	25.2	155	6	Sequence 57130, A
22	391	24.3	210	6	Sequence 55854, A
23	383	23.8	167	6	Sequence 40682, A
24	378	23.5	354	6	Sequence 36549, A
25	367	22.8	204	6	Sequence 57707, A
26	366.5	22.8	169	6	Sequence 45263, A
					Sequence 57507, A
					Sequence 40120, A

27	359.5	22.4	216	6	US-10-767-701-42579	Sequence 42579, A
28	355.5	22.1	178	6	US-10-767-701-32700	Sequence 32700, A
29	351	21.8	150	6	US-10-767-701-34317	Sequence 34317, A
30	349	21.7	341	6	US-10-767-701-45520	Sequence 45520, A
31	346.5	21.5	239	6	US-10-767-701-51756	Sequence 51756, A
32	339	21.1	153	6	US-10-767-701-61014	Sequence 61014, A
33	337.5	21.0	157	6	US-10-767-701-35363	Sequence 35363, A
34	333	20.7	213	6	US-10-767-701-45881	Sequence 45881, A
35	325.5	20.2	144	6	US-10-767-701-36220	Sequence 36220, A
36	324	20.1	144	6	US-10-767-701-57979	Sequence 57979, A
37	318.5	19.8	151	6	US-10-767-701-38693	Sequence 38693, A
38	313.5	19.5	198	6	US-10-767-701-45667	Sequence 45667, A
39	306.5	19.1	230	6	US-10-767-701-43621	Sequence 43621, A
40	305.5	19.0	186	6	US-10-767-701-51813	Sequence 51813, A
41	303	19.0	159	6	US-10-767-701-32166	Sequence 32166, A
42	296.5	18.4	147	6	US-10-767-701-61688	Sequence 61688, A
43	292	18.2	158	6	US-10-767-701-60922	Sequence 60922, A
44	290.5	18.1	138	6	US-10-767-701-38689	Sequence 38689, A
45	288.5	17.9	143	6	US-10-767-701-43202	Sequence 43202, A

ALIGNMENTS

RESULT 1

US-10-759-602-2

Sequence 2, Application US/10759602

GENERAL INFORMATION:

APPLICANT: Ainley, Michael

Armstrong, Katherine

Belmar, Scott

Folkerts, Otto

Hopkins, Nicole

Menke, Michael A.

Pareddy, Dayakar

Petolino, Joseph P.

Smith, Kelley

Woosley, Aaron

TITLE OF INVENTION: Regulatory Sequences for Transgenic Plants

NUMBER OF SEQUENCES: 59

CORRESPONDENCE ADDRESS:

ADDRESSEE: DowAgroSciences LLC

STREET: 9330 Zionsville Road

CITY: Indianapolis

STATE: Indiana

COUNTRY: USA

ZIP: 46268

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10759,602

FILING DATE: 16-Jan-2004

CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Kraus, Eric J

TELEPHONE: 317 337 5110

TELEFAX: 317 337 4847

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 333 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Query Match: 47.8%; Score 768.5; DB 6; Length 333;  
Best Local Similarity 49.7%; Pred. No. 2.7e-68;  
Matches 152; Conservative 52; Mismatches 99; Indels 3; Gaps 2;

NUMBER OF SEQ ID NOS: 63128  
 SEQ ID NO 46170  
 LENGTH: 357  
 TYPE: PRT  
 ORGANISM: Sorghum bicolor  
 FEATURE:  
 OTHER INFORMATION: Clone ID: SORBI-28MAY03-C17434\_1.pep  
 US-10-767-701-46170

Query Match 38.9%; Score 625; DB 6; Length 357;  
 Best Local Similarity 42.7%; Pred. No. 4.8e-54;  
 Matches 134; Conservative 48; Mismatches 116; Indels 16; Gaps 6;

QY 7 FYDNCSPVNSIVRTIVNELRSDPRTVSCADLLTIAAQQSVTLAGGSPWVPLGRD 126  
 DB 29 FYNKTPSAESIVQCTVAAAGNSGVAPAIRMHFDCFVRGCDGSLIDSTANTAEK 88  
 QY 67 DAFGNANSARGFPVIDRMKAAYESACPRTVSCADLLTIAAQQSVTLAGGSPWVPLGRD 126  
 DB 89 DSPANNFSLRFDDVDRAKASLEAQCFWVSCADLLTIAAQQSVTLAGGSPWVPLGRD 148  
 QY 127 SLQAFLLDANLNPAPFTTLPQKDSFRNVLNRSSDLVALSGGHTFGKQCRF----- 180  
 DB 149 GRVSNATQATNNLPPPPFFENATQVDRFASKNLT-LEDVWLSGAHTLGWSSGSPAGPAN 207  
 QY 181 IMDELYNFSNT--GLEPDTLNTVLOT-RGLCLPLNGNL---SALVDFDLRTPTTFDNKYY 235  
 DB 208 LGDELYNFSADGI-DPALS KAVAFLLKSI CPNSOFFNTTFMDIITPDKFNKYY 266  
 QY 236 VNLEEQGLIQSDQELFSSPDATDTIPLVRSPANSTQTFNFAVEMDRMGNITPLTGTO 295  
 DB 267 VGLTNLGLFESDAALLTNAIMK--ALVDSFVRNEITWKEKFAKSMVKMKIEVLGTQ 323  
 QY 295 GQIRLNCVWNSNS 309  
 DB 324 GEIRNCEVINPAS 337

RESULT 8  
 US-10-767-701-46411

Sequence 46411, Application US/10767701

APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement

FILE REFERENCE: 38-21(53535)B

CURRENT APPLICATION NUMBER: US/10/767,701

CURRENT FILING DATE: 2004-01-29

NUMBER OF SEQ ID NOS: 63128

SEQ ID NO 46411

LENGTH: 377

TYPE: PRT

ORGANISM: Sorghum bicolor

FEATURE:

OTHER INFORMATION: Clone ID: SORBI-28MAY03-C171286\_1.pep

US-10-767-701-46411

Query Match 38.0%; Score 611; DB 6; Length 377;  
 Best Local Similarity 43.3%; Pred. No. 1.2e-52;  
 Matches 133; Conservative 55; Mismatches 103; Indels 16; Gaps 7;

QY 7 FYDNCSPVNSIVRTIVNELRSDPRTVSCADLLTIAAQQSVTLAGGSPWVPLGR 125  
 DB 56 FYKSCPKAESIVKEFLSSAVQNVGLAALIRVHFDCFVQGCDSVLIDSTPTQPSQ 115  
 QY 67 DAFGNAN-SARGFPVIDRMKAAYESACPRTVSCADLLTIAAQQSVTLAGGSPWVPLGR 125  
 DB 116 LSPPNLTPRAFAKINDIRALEQACRVSCADLLTIAAQQSVTLAGGSPWVPLGR 175  
 QY 126 DLOAFLLDIA-NANLNPAPFTTLPQKDSFRNVLNRSSDLVALSGGHTFGKQCRFIMDR 184

DB 176 DGLAASNAVAALALPSTPTVTILLFLSKINLD-VTDLVALSGHTVGAHCS----- 229  
 QY 185 LYNFSNGLP--DPTLNTVLOTLAGLCELNGLSALVDFDLRTPTTFDNKYYVLEEQK 242  
 DB 230 --SFNRLFPPTQDPTLNKFFAGQLYGTCTPTTNTVTVN-DIRTPNTEFNKYYVDLLNRQ 286  
 QY 243 GLIQSDQELFSSPDATDTIPLVRSPANSTQTFNFAVEMDRMGNITPLTGTOGQIRLNC 302  
 DB 287 GLFTSDQDLLEN--ATTRPIVTKFAVDQNAEEFEQFVYVWVGQINVLITGSGQGVFRANC 343  
 QY 303 RVVNSNS 309  
 DB 344 SARNAGA 350

RESULT 9

US-10-767-701-44615

Sequence 44615, Application US/10767701

GENERAL INFORMATION:

APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement

FILE REFERENCE: 38-21(53535)B

CURRENT APPLICATION NUMBER: US/10/767,701

CURRENT FILING DATE: 2004-01-29

NUMBER OF SEQ ID NOS: 63128

SEQ ID NO 44615

LENGTH: 315

TYPE: PRT

ORGANISM: Sorghum bicolor

FEATURE:

NAME/KEY: unsure

LOCATION: (1) (315)

OTHER INFORMATION: unsure at all Xaa locations

FEATURE:

OTHER INFORMATION: Clone ID: SORBI-28MAY03-C56139\_1.pep

US-10-767-701-44615

Query Match 35.9%; Score 577.5; DB 6; Length 315;

Best Local Similarity 45.8%; Pred. No. 2e-49;

Matches 119; Conservative 41; Mismatches 87; Indels 13; Gaps 5;

QY 2 QLTPFTYDNCSPVNSIVRTIVNELRSDPRTVSCADLLTIAAQQSVTLAGGSPWVPLGR 61

DB 39 QLSSTFYDTCPRNALSTIKSGVDAAVMQEAETGASLRMHFDCFVHCGDGLVLLDTSQ 98

QY 62 ERTEKDAFGNANSARGFPVIDRMKAAYESACPRTVSCADLLTIAAQQSVTLAGGSPWVPLGR 121

DB 99 ---EQSSFPNKGSLRRFEDV-DSIKAQVEAVCPGVVSCADLLTIAAQQSVTLAGGSPWVPLGR 155

QY 122 LGRRLSLQAFLLDANLNPAPFTTLPQKDSFRNVLNRSSDLVALSGGHTFGKQCRFI 181

DB 156 LGRDSTASF-PSETTDLPTSSLOQLLSLFKNKLD-ATQMVVALSGAHTIGQAQCSNF 213

QY 182 MDRLYNFSNTGLPDTPTNTVLOTLAGLCELNGLSALVDFDLRTPTTFDNKYYVLEEQ 241

DB 214 NDHIYN-----DTNIDRAAFATSLQANCEASSTS-LAPLDTMTPTTFDNKYYVLEEQ 265

QY 242 KGLIQSDQELFSSPDATDTI 261

DB 266 KGLHSDQELFNNGSTDSTV 285

RESULT 10

US-10-767-701-43455

Sequence 43455, Application US/10767701

GENERAL INFORMATION:

APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement

; FILE REFERENCE: 38-21(53535)B

; CURRENT APPLICATION NUMBER: US/10/767,701

; CURRENT FILING DATE: 2004-01-29

; NUMBER OF SEQ ID NOS: 63128

; SEQ ID NO 43455

; LENGTH: 357

; TYPE: PRT

; ORGANISM: Sorghum bicolor

; FEATURE:

; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C35094\_1.pep

US-10-767-701-43455

Query Match 33.9%; Score 544.5; DB 6; Length 357;

Best Local Similarity 40.8%; Pred. No. 4.4e-46;

Matches 127; Conservative 38; Mismatches 135; Indels 11; Gaps 6;

QY 2 QLTPTFYDMSCPNVSNIVRDTIVNELRSDPRIAASILRLHFHDFCVNGCDASILLD-NTT 60

Db 34 QLKVGFFNYTTCPNRAELVQRVWTAAPANNNGVAPGLIRLHFHDFCVNGCDASVLLSNPA 93

QY 61 SEPTKDAFNGANSARGFPVIDRYKAAVESACPTVSCADLLTT-NAQGSVTLAGPSWRV 120

Db 94 GGNTERGSRANPSLRGDFVIDAATAAVERSCPTVSCADIVAFAPARDSVNLTKLFYQV 153

QY 121 PLGRDSLOAFDLANANLPAPFTLPOLKDSFRNVGLNRSSDLVALSGGTECKNOCRF 180

Db 154 PAGRRDGRVSNTEADNLLG2PSTAQVILGFRKNTL-VEDEVLSGSHSTIGRSHCAS 212

QY 181 INDRLYNFSNTGLPDPZTNTYTLQTLRGLCPN-GNLSA-LVPDFDLTPTTFIDNKYYVNL 238

Db 213 FLAT----NRRRLADGCTISAAYQALLCALCPSPGQDENTTEIDVSTPAVLDNNYKLL 268

QY 239 EEQKGLIQSQDELFSPPDADTTPLVRSFANSTQTFNFAVEAMDRMGNITELTG 293

Db 269 PMLGLHFSDDQIRN----ATLAPLANAFAPADTLWKQKFAAAMVWKGNDIVKGTITDEI 298

QY 299 RLKCRVNVNS 309

Db 326 RLKCSVVNPSS 336

RESULT 11

US-10-767-701-44200

; Sequence 44200, Application US/10767701

; GENERAL INFORMATION:

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement

; FILE REFERENCE: 38-21(53535)B

; CURRENT APPLICATION NUMBER: US/10/767,701

; CURRENT FILING DATE: 2004-01-29

; NUMBER OF SEQ ID NOS: 63128

; SEQ ID NO 44200

; LENGTH: 338

; TYPE: PRT

; ORGANISM: Sorghum bicolor

; FEATURE:

; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C46704\_1.pep

US-10-767-701-44200

Query Match

Best Local Similarity 39.9%; Score 519.5; DB 6; Length 338;

Matches 125; Conservative 41; Mismatches 128; Indels 19; Gaps 7;

QY 2 QLTPTFYDMSCPNVSNIVRDTIVNELRSDPRIAASILRLHFHDFCVNGCDASILLDNTTS 61

Db 37 QLOGACYNATCPAESLIETIVHAAVRKAGNGPGLIRLFFHDFCVNGCDASVLLDDPTG 96

QY 62 F---RT-EKDAFGNANSARGFPVIDRMKAAVESACPTVSCADLLTIAAQGSVTLAGPS 117

Db 97 TPGNRTVTKTSQPNFSLRGFSVINRAKRVVRRCPGTVSCADIVAFAPARDARIMGGR 156

QY 118 WRVPLGRDLSLOAFDLANANLPAPFTLPOLKDSFRNVGLNRSSDLVALSGGTECKNQ 177

Db 157 FAMPGRDLGVRVNASAEATANLPAPSENLTLQALFAASKNLT-ADDELVTLSGAHSIGRSH 215

QY 178 CRFIMDRLYNFSNTGL----PDPTLNTYTLQTLRGLCP-LNGNLSALVDFELTPTTFIDNK 233

Db 216 CS------SFANRLYPOLDATLNVTLAARAKCPAPGKGRVVTLDFTPLQLDNLQ 268

QY 234 YVNLLEQKGLIQSQDELFSPPDADTTPLVRSFANSTQTFNFAVEAMDRMGNITELTG 293

Db 269 YVSNVATHEVVGSDQAL----GDRNDCAALVALYANRKINSQKFAAMVWKGNS-EVLITG 325

QY 294 TQGOIRLNCRVVN 306

Db 326 PPGEVRLKCNKVN 338

RESULT 12

US-10-767-701-42477

; Sequence 42477, Application US/10767701

; GENERAL INFORMATION:

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement

; FILE REFERENCE: 38-21(53535)B

; CURRENT APPLICATION NUMBER: US/10/767,701

; CURRENT FILING DATE: 2004-01-29

; NUMBER OF SEQ ID NOS: 63128

; SEQ ID NO 42477

; LENGTH: 209

; TYPE: PRT

; ORGANISM: Sorghum bicolor

; FEATURE:

; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C52408\_1.pep

US-10-767-701-42477

Query Match

Best Local Similarity 31.3%; Score 503.5; DB 6; Length 209;

Matches 96; Conservative 31; Mismatches 48; Indels 1; Gaps 1;

QY 3 LTTFTFYDMSCPNVSNIVRDTIVNELRSDPRIAASILRLHFHDFCVNGCDASILLDNTTSF 62

Db 31 LFPQFYDHSCKPAKEIVQSIQAQAKETRMASLVRHLHFHDFCVKGCASVLLDNSSSI 90

QY 63 RTEKDAFGNANSARGFPVIDRMKAAVESACPTVSCADLLTIAAQGSVTLAGPSWRVPL 122

Db 91 VSEKSNPNRNSLRGFEVVDQIKAALEACPTVSCADIIAALARDSTSLVGGPYWDVPL 150

QY 123 GRDLSLOAFDLANANLPAPFTLPOLKDSFRNVGLNRSSDLVALSGGTECKNQ 178

Db 151 GRDLSLGSITQSGNSNDIPAPNTLPTITIKFKEQGLN-VDDVVALSGGTHIRMSRC 205

RESULT 13

US-10-767-701-44634

; Sequence 44634, Application US/10767701

; GENERAL INFORMATION:

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement

; FILE REFERENCE: 38-21(53535)B

; CURRENT APPLICATION NUMBER: US/10/767,701

; CURRENT FILING DATE: 2004-01-29

; NUMBER OF SEQ ID NOS: 63128

; SEQ ID NO 44634

; LENGTH: 341

; TYPE: PRT

```
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C7853_1.pep
US-10-767-701-44634

Query Match      30.5%; Score 490.5; DB 6; Length 341;
Best Local Similarity 37.8%; Pred. No. 9e-41;
Matches 115; Conservative 48; Mismatches 134; Indels 9; Gaps 5;

QY 2 QLTPTFYNSCPNWSNIVRDTIVNELRSDPRIAASILRLHFHDFCVNGCDASILLNNTTS 61
DB 36 RLSPNYRSCREVERIVSDVIAAKGRANPSTRAAGTILRLFFHDFCVNGCDASIVNPLSS 95
QY 62 FTEKDAFGNANSARG--FVIDRMKAAVESACPRVSCADLLTIAAQSVTLAGGPSWR 119
DB 96 TAAPRAABAINSLFGDAFDAVARAKAALESACPGVSCADVLAAARDLVLGLGPRPP 155
QY 120 VPLGRDLSLOAFDLANANLPAFFFTLPQLKDSFRVYGLNRSSDIVALSGCHTFGKNQCR 179
DB 156 VALGRDARSARCEVGNLPRTNMSARAVRLFAKGGLS-ADEMVALAGAHVGFSHCA 214
QY 180 FNDRIYNFSNTGLP--DPTLNTTYLQTLRGLOP-LNGNLSALVCFDLRPTTFIDNKYYV 236
DB 215 EFAHRIYGYRGAGAAGHDPLNPEFARALQRSAGYRDTPTWISFENDIVTPSEFDEAYK 274
QY 237 NLEBKGLLQSQDELFSPPDATDTIPLVRSFANSTQTFNFAVEMDRMGN-TIPLGTGQ 296
DB 275 NLZPHGLGLASDAAIWEYP---PTRVFAGRYGANSKRTAFEDFAAAMQRLGAVGKTRGQ 331
QY 297 QIRLNC 302
DB 332 VVRRRC 337

RESULT 14
US-10-767-701-35185
; Sequence 35185, Application US/10767701
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 35185
; LENGTH: 251
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C61249_1.pep
US-10-767-701-35185

Query Match      30.4%; Score 489; DB 6; Length 251;
Best Local Similarity 44.5%; Pred. No. 8.2e-41;
Matches 118; Conservative 29; Mismatches 104; Indels 14; Gaps 6;

QY 43 HDCFVNGCDASILLNNTTSFRTEKDAFGNANSARGPFPVIDRMKAAVESACPRVSCADLL 102
DB 1 HDCFVNGCDASILLNNTTSFRTEKDAFGNANSARGPFPVIDRMKAAVESACPRVSCADVL 59
QY 103 TIAAQSVTLAGGPSWRVPLGRDLSLOAFDLANANLPAFFFTLPQLKDSFRVYGLNRSS 162
DB 60 ALAARDAVLLARGPYVPLGRDGRSVDSTFTLPPFFNTSLIKLFSGHGFT-VQ 118
QY 163 DLVALSGCHTFGKNQCRFTIMDRLYNFSNTGLPDPILTNTTYLQTLRGCLFNGNLSALVDF 222
DB 119 DLVALSGCHTGLIAHCG-----NFKARLAETDLDALGSSSLGATCAANGDDGA-APP 170
QY 223 DLRTPTFDNKYYVNLKQGLIQSQDELFSPPDATDTIPLVRSFANSTQTFNFAVEM 282
```

```
DB 171 D-RTSTRFDVYVYRELQMRRLSSDQTLFESP---ETKGIVNFMAMQAYFFVAFQQGM 226
QY 283 DRMGNTPLGTGQGLRLNCRVNS 307
DB 227 LKMGQLDLKEGDEGEIRHTCGVNS 251

RESULT 15
US-10-767-701-46344
; Sequence 46344, Application US/10767701
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 46344
; LENGTH: 347
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C9288_1.pep
US-10-767-701-46344

Query Match      29.1%; Score 468; DB 6; Length 347;
Best Local Similarity 35.8%; Pred. No. 1.6e-38;
Matches 116; Conservative 53; Mismatches 107; Indels 50; Gaps 9;

QY 2 QLTPTFYNSCPNWSNIVRDTIVNELRSDPRIAASILRLHFHDFCVNGCDASILLNNTTS 61
DB 45 ELSVYFHVSCPLETIVSAVDAAALQONVRLTAGLLRVFFHDFCFQGCCDASILLD---- 100
QY 62 FTEKDAFGNANSARGFP-----VIDRMKAAVESACPRVSCADLLTIAAQSV 110
DB 101 -----NGERGLPPNWGLQBAVQIVEDIRAKVHAACGPTVSCADITVLAIRDAV 149
QY 111 TLAGGPSWRVPLGRDLSLOAFDLANAN---LPAPFFTLPLQKDSFRVYGLNRSSDLVA 166
DB 150 SLGGPSFTVPLGRDLSVAP---ASSNDFTLPPFTSTVDALLSAFASKNLSDDPDLVA 205
QY 167 LSGCHTFGKNQCRFTIMDRLYNFSNTGLPDPILTNTTYLQTLRGCLFNGNLSALVDFLRT 226
DB 206 LSGAHTWGRACS-----SFGDVAGP---ATDDITRCVTATCSAAGAGDTLRDLFLT 255
QY 227 PTIFDNKYV--NLEBKGL-LIQSQDELFSPPDATDTIPLVRSFANSTQTFNFAVEM 283
DB 256 PAVFDNLYFIETLTKXNKGWMLPSDQGLATDP---RTSLWLVQGFADNHNWFFDQFGTSMV 312
QY 284 RMGNITPLGTGQGLRLNCRVNS 309
DB 313 KMSQLKGPQNVGEIRNCLRENTNS 338

Search completed: April 6, 2004, 19:11:36
Job time : 7.97557 secs
```